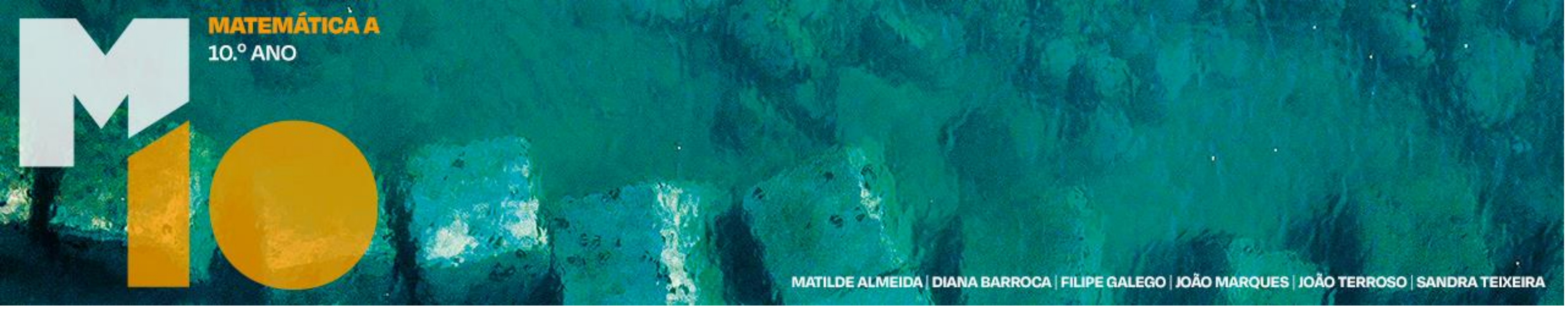


Tema 2 - Estatística

RECORDAR - Páginas 90 a 93

1. A frequência relativa da faixa etária "Idosos" é igual a:

- $Freq. Relativa$ Idosos $= 100\ \% - 11{,}48\ \% - 62{,}53\ \% = 25{,}99\ \%$
- A frequência absoluta da faixa etária "População com menos de 15 anos" é igual a:
- $Freq. Absoluta\ Jovem = 0{,}1148 \times 19\,348 = 2221{,}1504 \approx 2221$

A frequência absoluta da faixa etária "População em idade ativa" é igual a:

- $Freq. Absoluta\ idade\ ativa = \ 0{,}6253 \times 19\,348 = 12\,098{,}3044 \approx 12\,098$
- A frequência absoluta da faixa etária "Idosos" é igual a:
- $Frequência\ Absoluta\ Idosos = \ 0{,}2599 \times 19\,348 = 5028{,}5452 \approx 5029$

Podemos então completar a tabela da seguinte forma:

Residentes no Concelho da Mealhada, em 2021

Faixa etária	Frequência relativa (%)	Frequência absoluta
Jovens com menos de 15 anos	11,48%	**2221**
População em idade ativa (15 aos 64 anos)	62,53%	**12 098**
Idosos ($\geq 65\ anos$)	**25,99%**	**5029**
Total	**100,00%**	**19 348**

2.

2.1. Calculando a média de cada um dos conjuntos de dados, temos que:

- **<u>Conjunto A:</u>** $\bar{a} = \frac{19+20+17+20+10+17}{6} \approx 17{,}17$
- **<u>Conjunto B:</u>** $\bar{b} = \frac{1{,}64+1{,}75+1{,}68+1{,}90}{4} \approx 1{,}74$
- **<u>Conjunto C:</u>** $\bar{c} = \frac{1+0+2+0+0+3+5}{7} \approx 1{,}57$

2.2. O conjunto A pode representar, por exemplo, as idades de um grupo de amigos.
O conjunto B pode representar a altura, em metros, de um conjunto de pessoas.
O conjunto C pode representar o número de irmãos de 7 alunos de uma turma.

3.

3.1. Os números de músicas nas listas do Joaquim variam entre 9 e 15. Podemos então organizar os dados na tabela de frequências absolutas seguinte.

Número de músicas das várias listas do Joaquim

Número de músicas	9	10	11	12	13	14	15
Frequência absoluta	**2**	**5**	**3**	**1**	**2**	**2**	**1**

3.2. Tendo em conta os resultados organizados na tabela de frequências absolutas:

$$\bar{x} = \frac{9 \times 2 + 10 \times 5 + 11 \times 3 + 12 \times 1 + 13 \times 2 + 14 \times 2 + 15 \times 1}{16} = \frac{182}{16} = 11{,}375$$

A média de músicas nas listas do Joaquim é 11,375.

3.3. A moda deste conjunto de dados é 10 músicas.

4. Tendo em conta que a média de chegadas é de 1122 milhares, então:

$$\bar{x} = \frac{a + 300 + 2900 + 1500 + 262 + 1000}{6} = 1122 \Leftrightarrow$$
$$\Leftrightarrow a + 5962 = 6 \times 1122 \Leftrightarrow$$
$$\Leftrightarrow a = 6732 - 5962 \Leftrightarrow$$
$$\Leftrightarrow a = 770$$

5.

5.1. A expressão representa o cálculo da média das idades dos alunos da turma.

$$\frac{2 \times 15 + 11 \times 16 + 10 \times 17 + 2 \times 18}{25} = 16{,}48$$

5.2. A média das idades da turma A é igual a 16,48. A média das idades dos alunos da turma B do 10.º ano é 16,1.

Não podemos tirar essa conclusão. Por exemplo, no caso da turma 10.º B, ter 18 alunos com 16 anos e 2 alunos com 17, a idade média do 10.º B seria de 16,1 anos. No entanto, em termos comparativos, a turma 10.º A tem mais alunos com idades superiores a 16 anos do que a turma 10.º B deste exemplo.

6.

6.1. Temos que:

- $N.º\ de\ alunos = 2 + 7 + 12 + 4 = 25$

Realizaram a questão de aula 25 alunos.

6.2. Nesta questão de aula, apresentaram níveis positivos (classificação superior ou igual a 10):

$$N.º\ de\ alunos\ com\ positiva = 12 + 4 = 16$$

Logo, a percentagem de níveis positivos foi:

$$Percentagem\ de\ níveis\ positivos = \frac{16}{25} = 0{,}64 = 64\ \%$$

A percentagem de alunos que apresentaram níveis positivos é 64 %.

6.3. Resposta de comentário livre.

7. Com os dados da tabela podemos determinar:

- $a = 389 - (45 + 74 + 16 + 115) = 139$
- $b = \frac{74}{389} \approx 19{,}0\%$
- $c = \frac{16}{389} \approx 4{,}1\%$

Concluímos que $a = 139$, $b \approx 19{,}0\%$ e $c \approx 4{,}1\%$.

8.

8.1. Tendo em conta o gráfico apresentado, podemos construir a tabela de frequências seguinte.

Idade dos trabalhadores

Idade	Número de trabalhadores
[20,25[	**30**
[25,30[	**60**
[30,35[	**80**
[35,40[	75
[40,45[	**50**
[45,50[	**60**

8.2. Como não conhecemos os valores das idades dos trabalhadores da empresa, não conseguimos determinar o valor exato da média das idades destes trabalhadores.

8.3. Tem-se que:

- $Número\ de\ trabalhadores\ da\ empresa\ = 30 + 60 + 80 + 75 + 50 + 60 = 355$

Já o número de trabalhadores com idade superior ou igual a 35 anos é:

- $N.º\ de\ trabalhadores = 75 + 50 + 60 = 185$

Logo:

- $Percentagem\ de\ trabalhadores\ com\ idade\ superior\ ou\ igual\ a\ 35 = \frac{185}{355} \approx 52{,}1\ \%$

A percentagem de trabalhadores com idade superior ou igual a 35 anos é de, aproximadamente, 52,1%.

9.

9.1. Esquematizando o conjunto de dados por ordem crescente, tem-se que:

Ordem	Tempos de Leitura Ordenados	
1	2	
2	2	
3	2	
4	2	
5	3	
6	3	Valores centrais da primeira metade
7	3	
8	3	
9	4	
10	4	
11	4	
12	4	Valores centrais
13	5	
14	6	
15	6	
16	9	
17	9	
18	9	Valores centrais da segunda metade
19	10	
20	11	
21	11	
22	13	
23	20	
24	24	

Assim, podemos determinar o valor da mediana e dos quartis da seguinte forma:

$$Q_1 = \frac{3+3}{2} = 3 \qquad Q_2 = mediana = \frac{4+5}{2} = 4{,}5 \qquad Q_3 = \frac{9+10}{2} = 9{,}5$$

Concluímos que o 1.º quartil é 3, a mediana é 4,5 e o 3.º quartil é 9,5.

9.2. Sabemos que $Q_1 = 3$ e $Q_3 = 9{,}5$.

A amplitude interquartil é:

$$Amplitude\ Interquartis = Q_3 - Q_1 = 9{,}5 - 3 = 6{,}5$$

Neste conjunto de dados, a amplitude interquartil é 6,5.

9.3. Temos que $min. = 2$ e $máx. = 24$.

A amplitude deste conjunto de dados é:

$$Amplitude = máx. - min. = 24 - 2 = 22$$

A amplitude deste conjunto de dados é 22.

2.1. Problema estatístico
Tarefa 1 - Página 94

Resposta ao cuidado do aluno.

Aplicar - Página 95
1.
1.1. Fazendo os respetivos cálculos, temos que:

- $N.º\ bolas\ de\ berlim = 2500 + 2450 + 2480 + 2600 + 2700 + 3000 + 50 = 16\,230$
- $N.º\ pães\ da\ avó = 1200 + 1100 + 1150 + 1300 + 1350 + 1600 + 200 = 7900$
- $N.º\ croissants = 220 + 230 + 210 + 250 + 260 + 600 + 120 = 1890$

Durante a semana, o sr. Manuel vendeu, ao todo, 16 230 bolas, 7900 pães da avó e 1890 croissants.
O produto mais popular da padaria do sr. Manuel são as bolas de Berlim.

1.2. Fazendo a apuração das vendas totais registadas ao longo das semanas, temos que:

- $Vendas\ 2.ª\ feira = 2500 + 1200 + 220 = 3920$
- $Vendas\ 3.ª\ feira = 2450 + 1100 + 230 = 3780$
- $Vendas\ 4.ª\ feira = 2480 + 1150 + 210 = 3840$
- $Vendas\ 5.ª\ feira = 2600 + 1300 + 250 = 4150$
- $Vendas\ 6.ª\ feira = 2700 + 1350 + 260 = 4310$
- $Vendas\ Sábado = 3000 + 1600 + 600 = 5200$
- $Vendas\ Domingo = 500 + 200 + 120 = 820$

O dia da semana em que o sr. Manuel teve vendas mais baixas foi no domingo, pelo que deverá encerrar ao domingo.

Aplicar - Página 97

2.

2.1. A população em estudo é o conjunto dos alunos do ensino secundário da escola da Maria.

2.2. Esta amostra tem dimensão igual a 50.

2.3.

a) A variável estatística da tabela A é a idade, em anos, dos alunos.
b) A variável estatística da tabela B é o número de dias que cada um dos alunos da amostra pratica deporto, por semana.
c) A variável estatística da tabela C é a massa, em quilogramas, de cada um dos alunos da amostra.

Tarefa 2 - Página 98

Resposta de pesquisa livre. Ao cuidado dos alunos.

Aplicar - Página 99

3.

3.1. Podem ser colocadas, por exemplo, as questões seguintes.

- Quantas refeições diárias faz normalmente?
- Tendo em conta os seus hábitos de consumos, considera que a sua dieta é rica em hidratos de carbono, proteínas ou gorduras?
- Costuma beber 1 copo de água diariamente?
- Costuma comer salada nas duas refeições principais diárias (almoço e jantar)?
- Quantas peças de fruta costuma comer diariamente?
- Com que frequência usa especiarias na elaboração das suas refeições? De que tipo?

3.2. Para a produção e aquisição de dados deve ser inquirida toda a população, pois o número de elementos da população é pequeno.

3.3. Para organizar e representar os dados recolhidos, deve-se: realizar a contagem, o agrupamento e sistematização dos dados; elaborar tabelas de frequências e gráficos; determinar, por exemplo, a moda, a média e a mediana dos conjuntos de dados, se possível.

3.4. Para interpretação dos dados, dever-se-ão analisar as tabelas e os gráficos obtidos, bem como as medidas estatísticas calculadas. De seguida, dever-se-ão registar as conclusões, tendo em conta o contexto da situação e a problemática associada.

Tarefa 3 - Página 100

1. A população em estudo é constituída pelo conjunto de eleitores portugueses residentes em Portugal no ano 2023.

2. São 1006 os elementos da amostra e foram selecionados aleatoriamente a partir de uma lista de números de telemóvel, também ela gerada de forma aleatória.

3. Neste inquérito, foi feita uma sondagem, porque foi selecionada uma amostra de entre os elementos da população.

4. Neste inquérito, a ficha técnica esclarece-nos que todos os inquiridos foram informados do objetivo do estudo. Informa-nos ainda sobre as caraterísticas da amostra relativamente ao género, distribuição geográfica, escalões etários e qual a taxa de respostas. Somos ainda esclarecidos sobre a margem de erro máxima e o nível de confiança.

5. Ao cuidado do aluno.

Aplicar - Página 101

4. Ao cuidado do aluno. Recomendam-se, por exemplo, os seguintes sites para pesquisa:

- https://observador.pt/seccao/sociedade/sondagens/
- https://www.publico.pt/sondagem
- https://www.publico.pt/censos

5. Tendo em conta que se trata de um estudo estatístico aplicado à população de uma cidade, a realização de uma sondagem é a melhor opção.

Aplicar - Página 102

6. Ao cuidado do aluno.

Aplicar - Página 103

7.

7.1. Existe risco de enviesamento da amostra, pois os alunos que frequentam a biblioteca podem ter hábitos de leitura diferentes da generalidade dos alunos da escola.

7.2. Existe risco de enviesamento, porque os alunos que frequentam o desporto escolar podem ter hábitos de prática de atividade física (por gostarem de desporto, por terem treinos regulares) diferentes da generalidade dos alunos da escola.

7.3. Existe risco de enviesamento. Os números de telefone fixo são selecionados aleatoriamente, o que poderá levar à seleção de eleitores não seguidores da ideologia política do partido A, o que levará a respostas tendenciosas ou não respostas.

Aplicar - Página 104

8. Dever-se-á numerar, de 1 a 200, todos os alunos do 10.º ano do Agrupamento de Escolas da Abraçada. Depois, procede-se a seleção aleatória, com recurso à tecnologia, de 50 números. Os números selecionados correspondem aos números dos alunos que irão integrar a amostra.

9. Começamos por determinar o número total de habitantes:

$$N.º\ de\ habitantes = 677 + 2370 + 813 + 976 + 3143 + 2824 + 1618 + 713 = 13\ 134$$

A amostra estratificada deve representar, de forma proporcional, cada uma das freguesias.
Assim, para calcularmos quantos representantes de cada freguesia irão pertencer à amostra, fazemos o seguinte.

- Freguesia de Carvalho: $\frac{677\times 260}{13\ 134}\approx 13{,}40$
- Freguesia de Figueira de Lorvão: $\frac{2370\times 260}{13\ 134}\approx 46{,}92$
- Freguesia de Friumes e Paradela: $\frac{813\times 260}{13\ 134}\approx 16{,}09$
- Freguesia de Oliveira do Mondego e Travanca do Mondego: $\frac{976\times 260}{13\ 134}\approx 19{,}32$
- Freguesia de Lorvão: $\frac{3143\times 260}{13\ 134}\approx 62{,}22$
- Freguesia de Penacova: $\frac{2824\times 260}{13\ 134}\approx 55{,}90$
- Freguesia de São Pedro de Alva e São Paio do Mondego: $\frac{1618\times 260}{13\ 134}\approx 32{,}03$
- Freguesia de Sazes do Mondego: $\frac{713\times 260}{13\ 134}\approx 14{,}11$

Se procedermos à adição dos valores arredondados às unidades, obtemos:
- $N.º\ de\ elementos = 13+47+16+19+62+56+32+14=259$

Falta um elemento, pelo que temos duas hipóteses: ou optamos por assumir uma amostra de 259 elementos; ou teremos de decidir a que freguesia atribuir mais um elemento.
Optando pela segunda hipótese, podemos, por exemplo, por atribuir esse elemento à freguesia que, tendo sido aproximada por defeito, tenha um valor mais perto da unidade seguinte. Neste caso, é a freguesia de Carvalho, que poderá ficar com 14 elementos na amostra.

Assim, dos 62 elementos da amostra, teremos: 14 da freguesia de Carvalho, 47 de Figueira de Lorvão, 16 de Friúmes e Paradela, 19 de Oliveira do Mondego e Travanca do Mondego, 62 de Lorvão, 56 de Penacova, 32 de são Pedro de Alva e São Paio de Mondego e 14 de Sazes do Lorvão.

Aplicar - **Página 105**

10.

10.1. No 10.º ano, frequentam esta escola:
- $N.º\ de\ alunos\ do\ 10.º\ ano = 27+24+23+26+20=120$

Nesta escola, existem 120 alunos a frequentar o 10.º ano.

10.2. A população em estudo é o conjunto dos alunos que frequentam o 10.º ano desta escola.

10.3. A amostra deverá incluir 20% da população. Assim:
- $Dimensão\ da\ amostra = \frac{20}{100}\times 120 = 24$

A amostra deverá incluir 24 alunos do 10.º ano desta escola.

10.4. A amostra deverá ser selecionada tendo em conta a percentagem de alunos de cada turma do 10.º ano. Assim:
- $N.º\ de\ alunos\ amostra\ 10.º\ A = \frac{27}{120}\times 24 = 5{,}4\approx 5$
- $N.º\ de\ alunos\ amostra\ 10.º\ B = \frac{24}{120}\times 24 = 4{,}8\approx 5$

- $N.º\ de\ alunos\ amostra\ 10.º\ C = \frac{23}{120} \times 24 = 4{,}6 \approx 5$
- $N.º\ de\ alunos\ amostra\ 10.º\ D = \frac{26}{120} \times 24 = 5{,}2 \approx 5$
- $N.º\ de\ alunos\ amostra\ 10.º\ E = \frac{20}{120} \times 24 = 4$

Concluímos que a amostra deverá incluir 5 alunos de cada uma das turmas A, B, C e D e 4 alunos da turma E.

Aplicar - Página 106

11. Podemos identificar, por exemplo, os seguintes termos estatísticos: sondagem, universo, amostra, margem de erro, intervalo de confiança.

2.2. Dados Univariados

Página 107

Tarefa 4

1. A população em estudo é o conjunto dos alunos da escola do Manuel nesse ano letivo.

2. Estão a ser estudadas as seguintes variáveis:

- ano de escolaridade;
- idade (em anos);
- ter (ou não) computador com ligação à *internet*;
- finalidade da utilização do computador;
- utilização de redes sociais;
- tempo diário de ligação à *internet* (em minutos).

3. As variáveis podem ser classificadas da seguinte forma:

- Ano de escolaridade é uma variável qualitativa ordinal.
- Idade (em anos) é uma variável quantitativa discreta.
- Ter (ou não) computador com ligação à *internet* é uma variável qualitativa nominal.
- Finalidade da utilização do computador é uma variável qualitativa nominal.
- Utilização de redes sociais é uma variável qualitativa nominal.
- Tempo ligado à *internet* (em minutos) é uma variável quantitativa contínua.

4. Ao cuidado do aluno.

Aplicar - Página 109

12.

A. Variável quantitativa contínua.
B. Variável quantitativa contínua.
C. Variável quantitativa discreta.
D. Variável qualitativa ordinal.
E. Variável quantitativa contínua.
F. Variável quantitativa discreta.
G. Variável quantitativa contínua.
H. Variável qualitativa nominal.
I. Variável qualitativa nominal.

Tarefa 5 - Página 110

O número de alunos do professor Manuel é 100. Assim, podemos completar a tabela da seguinte forma:

Número de questões resolvidas pelos alunos

Número de questões resolvidas	Frequência absoluta	Frequência relativa	Frequência absoluta acumulada	Frequência relativa acumulada
0	10	0,1	10	0,1
2	12	0,12	22	0,22
4	8	0,08	30	0,3
6	15	0,15	45	0,45
8	13	0,13	58	0,58
9	5	0,05	63	0,63
10	7	0,07	70	0,7
12	11	0,11	81	0,81
14	3	0,03	84	0,84
16	9	0,09	93	0,93
17	3	0,03	96	0,96
20	4	0,04	100	1
Total	**100**	**1**		

Aplicar - Página 111

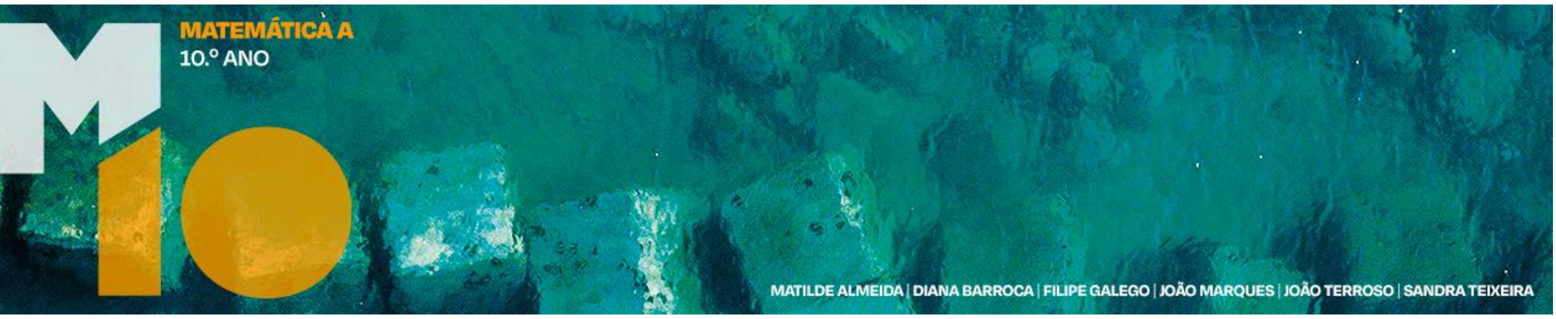
13.

13.1. A variável em estudo é o número de peixes pescados por cada um dos participantes. Esta variável é quantitativa discreta.

13.2. O número de peixes pescados pelos participantes varia entre 0 a 10.
Podemos organizar os dados recolhidos numa tabela de frequências semelhante à seguinte.

Número de peixes pescados pelos participantes no concurso do clube dos pescadores

N.º de Peixes	Frequência Absoluta	Frequência Absoluta Acumulada	Frequência Relativa	Frequência Relativa Acumulada
0	1	1	$\frac{1}{25}=0{,}04$	$0{,}04$
1	3	$1+3=4$	$\frac{3}{25}=0{,}12$	$0{,}04+0{,}12=0{,}16$
2	2	$4+2=6$	$\frac{2}{25}=0{,}08$	$0{,}16+0{,}08=0{,}24$
3	0	$6+0=6$	0	$0{,}24+0=0{,}24$
4	3	$6+3=9$	$\frac{3}{25}=0{,}12$	$0{,}24+0{,}12=0{,}36$
5	4	$9+4=13$	$\frac{4}{25}=0{,}16$	$0{,}36+0{,}16=0{,}52$
6	4	$13+4=17$	$\frac{4}{25}=0{,}16$	$0{,}52+0{,}16=0{,}68$
7	3	$17+3=20$	$\frac{3}{25}=0{,}12$	$0{,}68+0{,}12=0{,}8$
8	2	$20+2=22$	$\frac{2}{25}=0{,}08$	$0{,}8+0{,}08=0{,}88$
9	1	$22+1=23$	$\frac{1}{25}=0{,}04$	$0{,}88+0{,}04=0{,}92$
10	2	$23+2=25$	$\frac{2}{25}=0{,}08$	$0{,}92+0{,}08=1$
Total	25		1	



Aplicar - Página 112

14.

14.1. Na turma do Eduardo, a altura mínima é 145 cm e a máxima é 192 cm.

14.2. Tendo em conta que a altura mínima registada é de 145 cm e que cada classe a considerar tem de amplitude 10 unidades, podemos construir a tabela de frequências simples e acumuladas seguinte.

Alturas, em centímetros, dos alunos da turma do Eduardo

Altura (cm)	Frequência absoluta	Frequência relativa	Frequência absoluta acumulada	Frequência relativa acumulada
$[145; 155[$	4	$\frac{4}{24} \approx 0{,}17$	4	$0{,}17$
$[155; 165[$	6	$\frac{6}{24} = 0{,}25$	$4 + 6 = 10$	$0{,}17 + 0{,}25 = 0{,}42$
$[165; 175[$	2	$\frac{2}{24} \approx 0{,}08$	$10 + 2 = 12$	$0{,}42 + 0{,}08 = 0{,}5$
$[175; 185[$	8	$\frac{8}{24} \approx 0{,}33$	$12 + 8 = 20$	$0{,}5 + 0{,}33 = 0{,}83$
$[185; 195[$	4	$\frac{4}{24} \approx 0{,}17$	$20 + 4 = 24$	$0{,}83 + 0{,}17 = 1$
Total	24	1		

14.3. A altura do Eduardo está contida na classe $[175; 185[$.
Consultando a tabela construída, verificamos que esta classe regista uma frequência relativa de, aproximadamente, $0{,}33$.

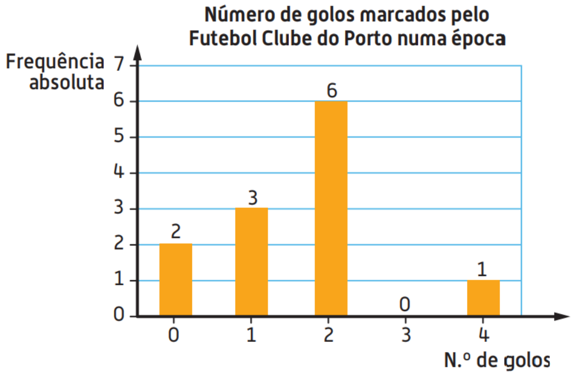
Aplicar - Página 113

15. Começamos por organizar os dados na tabela de frequências absolutas seguinte.

Número de golos marcados pelo Futebol Clube do Porto numa época

Número de golos	Frequência absoluta
0	2
1	3
2	6
3	0
4	1
Total	12

Com os dados da tabela, conseguimos construir o seguinte gráfico de barras.



Fonte: https://www.fcporto.pt/pt/futebol/fcporto

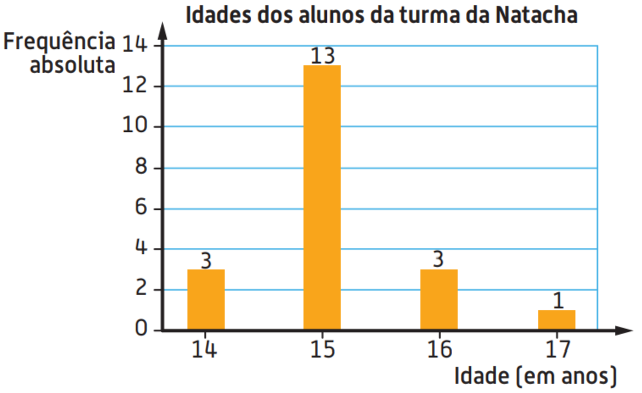
16.

16.1. Podemos determinar a frequência absoluta das idades adicionando o número de elementos do género feminino com o número de elementos do género masculino associados. Obtemos assim a tabela seguinte.

Idades dos alunos da turma da Natacha

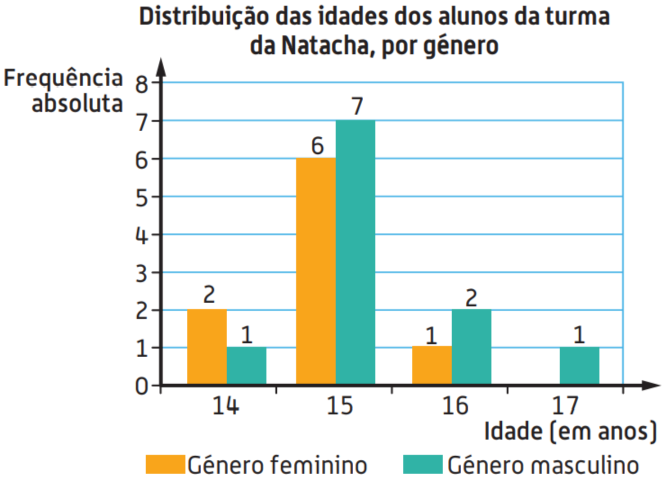
Idade	Frequência absoluta
14	3
15	13
16	3
17	1
Total	20

Com os dados recolhidos na tabela, construímos o gráfico de barras seguinte.



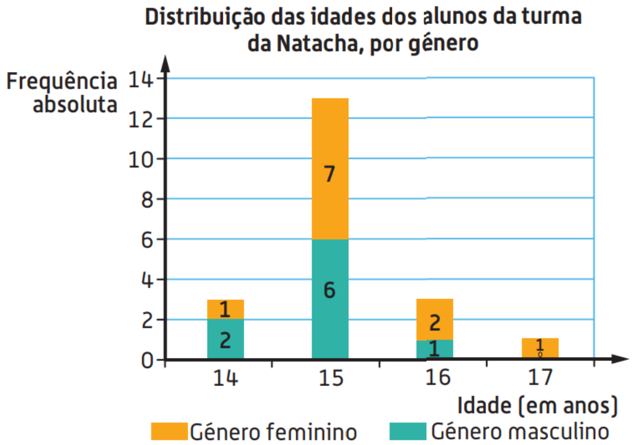
Fonte: Própria

16.2. A distribuição das idades pelos géneros masculino e feminino da turma da Natacha permite-nos representar o gráfico de barras justapostas seguinte.



Fonte: Própria

16.3. De modo semelhante, com a distribuição de idades pelos géneros da turma da Natacha representamos os dados no gráfico de barra sobrepostas seguinte.



Fonte: Própria

Tarefa 6 - Página 114

1. Segundo os dados representados nos dois histogramas, temos que:

Valor do IMC dos ouvintes do "A sua tarde na *OnOff*"

IMC	Frequência absoluta	Frequência absoluta acumulada
[14,18[	18	18
[18,22[	26	$18 + 26 = 44$
[22,26[	6	$44 + 6 = 50$
[26,30[	32	$50 + 32 = 82$
[30,34[	16	$82 + 16 = 98$
Total	98	

Valor do IMC dos ouvintes do "*OnOff night*"

IMC	Frequência absoluta	Frequência absoluta acumulada
[14,18[	11	11
[18,22[	$27 - 11 = 16$	27
[22,26[	$41 - 27 = 14$	41
[26,30[	$42 - 41 = 1$	42
[30,34[	$50 - 42 = 8$	50
Total	50	

Logo, conjugando os dados recolhidos nas emissões dos dois programas, podemos obter a seguinte tabela:

Valor do IMC no total de ouvintes

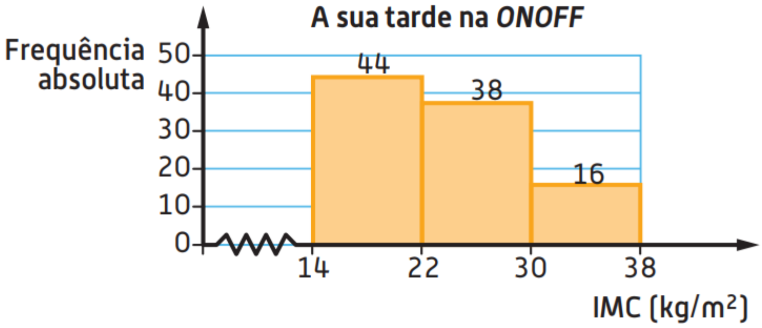
IMC	Frequência absoluta	Frequência absoluta acumulada
[14,18[	$11 + 18 = 29$	29
[18,22[	$26 + 16 = 42$	$29 + 42 = 71$
[22,26[	$14 + 6 = 20$	$71 + 20 = 91$
[26,30[	$32 + 1 = 33$	$91 + 33 = 124$
[30,34[	$16 + 8 = 24$	$124 + 24 = 148$
Total	148	

2. Com a alteração de classes, reorganizamos os dados tabela de frequências absolutas seguinte.

Valor do IMC dos ouvintes do "A sua tarde na *OnOff*"

IMC	Frequência absoluta
[14,22[	44
[22,30[	38
[30,38[	16
Total	98

O histograma associado a esta nova distribuição de dados é o seguinte.



Fonte: Rádio ONOFF

Analisando e comparando os dois histogramas, podemos verificar que aquele com classes de menor amplitude apresenta um maior nível de detalhe no que diz respeito à distribuição dos dados ao contrário do histograma cujas classes têm maior amplitude.

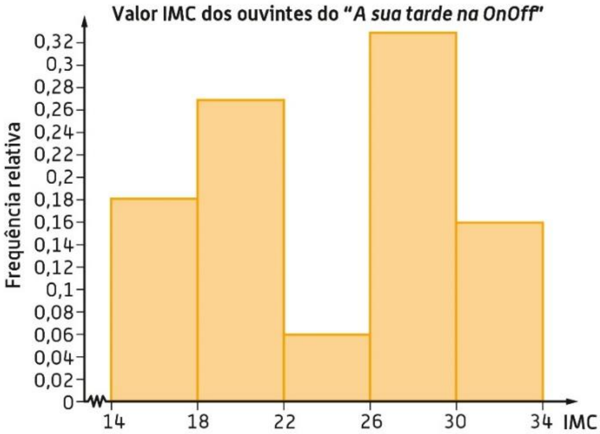
3. Comecemos por determinar a frequência relativa das classes iniciais dos ouvintes do programa "A sua tarde na OnOff".

Tem-se que:

Valor do IMC dos ouvintes do "A sua tarde na *OnOff*"

IMC	Frequência absoluta	Frequência Relativa
[14,18[	18	$\frac{18}{98} \approx 0{,}18$
[18,22[	26	$\frac{26}{98} \approx 0{,}27$
[22,26[	6	$\frac{6}{98} \approx 0{,}06$
[26,30[	32	$\frac{32}{98} \approx 0{,}33$
[30,34[	16	$\frac{16}{98} \approx 0{,}16$
Total	98	1

Fazendo a representação do histograma associado, obtemos a representação gráfica seguinte.



Fonte: Própria

Podemos então determinar a área do histograma construído:

$$Área_{histograma} = Área_{1.ª\ classe} + Área_{2.ª\ classe} + Área_{3.ª\ classe} + Área_{4.ª\ classe} + Área_{5.ª\ classe} =$$
$$= 4 \times 0{,}18 + 4 \times 0{,}27 + 4 \times 0{,}06 + 4 \times 0{,}33 + 4 \times 0{,}16 = 4$$

Podemos verificar que a área do histograma de frequência relativas simples coincide com a amplitude de cada uma das classes consideradas.

Aplicar - Página 115

17.

17.1. Para a construção da tabela de frequências, consideremos as classes seguintes.

$[40{,}70[\qquad [70{,}100[\qquad [100{,}130[\qquad [130{,}160[\qquad [160{,}190[\qquad [190{,}220[$

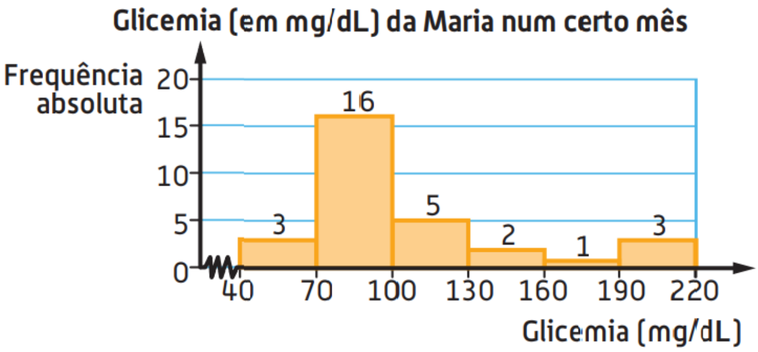
Organizando os dados, construímos a tabela seguinte.

Glicemia (em mg/dL) da Maria durante um mês

Glicemia em mg/dL	Frequência absoluta
[40,70[	3
[70,100[	16
[100,130[	5
[130,160[	2
[160,190[	1
[190,220[	3
Total	30

17.2.

a) Usando os dados organizados na tabela de frequências absolutas da alínea anterior, podemos representar o histograma seguinte.



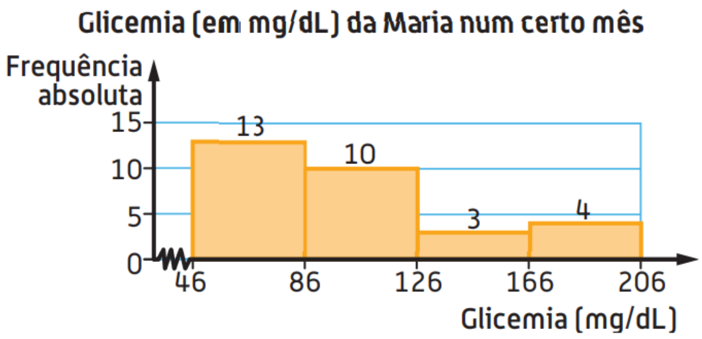
Fonte: Própria

b) Representamos os dados numa nova tabela de frequências absolutas.

Glicemia (em mg/dL) da Maria durante um mês

Glicemia em mg/dL	Frequência absoluta
[46,86[	13
[86,126[	10
[126,166[	3
[166,206[	4
Total	30

De seguida, construímos o histograma seguinte.



Fonte: Própria

Aplicar - Página 116

18. Calculando o número total de jornalistas, as frequências relativas simples e as frequências relativas acumuladas obtemos a tabela seguinte.

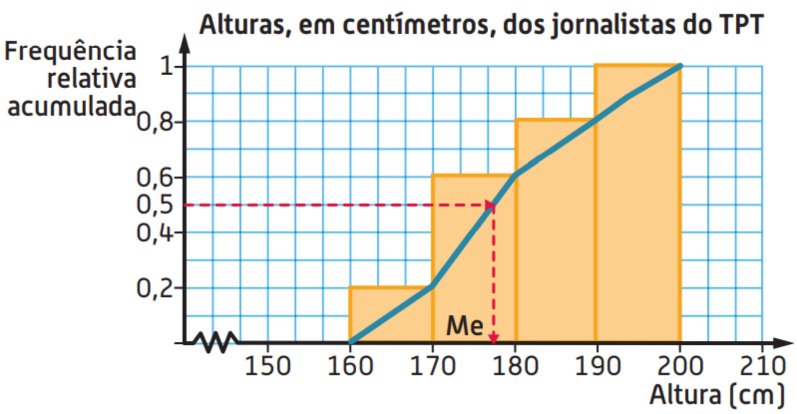
Alturas, em centímetros, dos jornalistas do TPT

Altura (em centímetros)	Número de jornalistas	Frequência relativa simples	Frequência relativa acumulada
[160,170[	4	$\frac{4}{20}=0{,}2$	0,2
[170,180[	8	$\frac{8}{20}=0{,}4$	$0{,}2+0{,}4=0{,}6$
[180,190[	6	$\frac{6}{20}=0{,}3$	$0{,}6+0{,}3=0{,}9$
[190,200[	2	$\frac{2}{20}=0{,}1$	$0{,}9+0{,}1=1$
Total	20	1	—

A classe mediana é a classe [170,180[.

A partir dos dados da tabela, construímos um histograma com as frequências relativas acumuladas e o polígono de frequências relativas acumuladas.

Depois, identificando o ponto do polígono de frequências relativas acumuladas que corresponde à frequência relativa acumulada igual a 0,5, podemos determinar, por processos geométricos, o valor aproximado da altura mediana, como se observa no histograma seguinte.



Fonte: Própria

Tarefa 7 - Página 117

1. Contando o número de entradas no diagrama de caule-e-folhas representado, podemos concluir que a Associação tem 32 sócios.

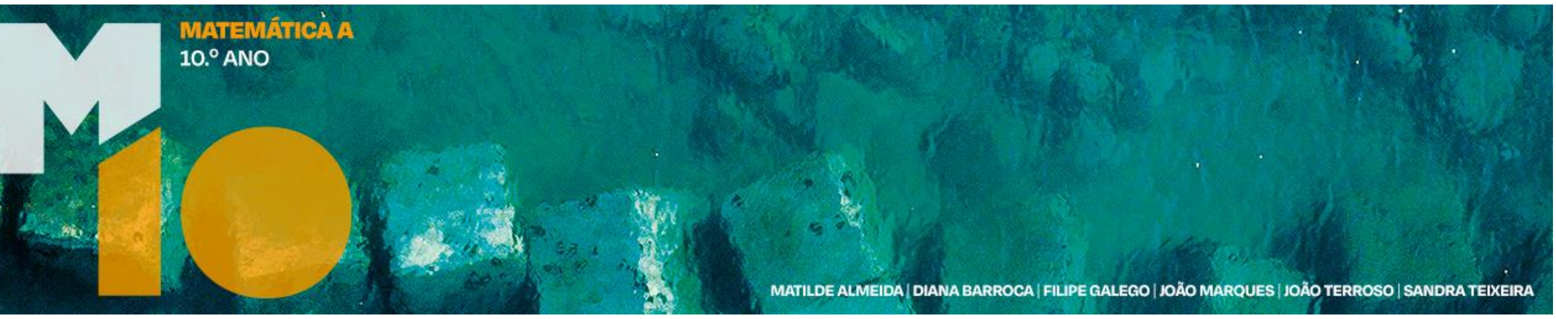
2. Os sócios fundadores serão os três sócios cujo tempo de associado é de 50 anos, ou seja, os que são sócios há mais tempo. Permanecem na Associação 3 sócios fundadores.

3. Existem 9 sócios que registam 30 ou mais anos de associação.

4. Com a representação dos dados no diagrama de caule-e-folhas, podemos ordenar, de forma crescente, os dados da forma seguinte.

1, 2, 3, 4, 4, 4, 10, 10, 10, 11, 11, 11, 11, 12, 13, 14, 15

21, 21, 22, 22, 23, 24, 30, 31, 32, 41, 48, 49, 50, 50, 50



Aplicar - Página 118

19.
19.1. Os registos do Francisco permitem construir o diagrama de caule-e-folhas seguinte.

Quantidade de mel, em quilogramas, em cada colmeia

0	4 5
1	1 1 2 6 7 8 9
2	0 0 0 2 9
3	2 5 5 8
4	1 2

0|4 significa 4 kg

Fonte: Própria

19.2. O conjunto de dados recolhidos pelo Francisco podem ser ordenadas da forma seguinte.

4, 5, 11, 11, 12, 16, 17, 18, 19, 20, 20, 20, 22, 29, 32, 35, 35, 38, 41, 42

20.
20.1. O rancho folclórico tem 26 membros.

20.2. Podemos concluir que sim, pois há tantos membros do género masculino como do género feminino.

20.3. Existem 6 membros do rancho com mais de 40 anos de idade (3 do género feminino e 3 do género masculino).

20.4. O rancho tem mais membros do sexo feminino do que membros do género masculino com mais de 50 anos (3 membros do género feminino e 2 membros do género masculino nesta faixa etária).

Tarefa 8 - Página 119

1. O gráfico de extremos e quartis permite concluir que o valor mínimo de praias é de 286.
2. O valor máximo de praias distinguidas com "Qualidade de Ouro" foi de 396.

3. Neste conjunto de dados, o valor mediano de praias distinguidas com a classificação "Qualidade de Ouro" foi 345,5.

4. Neste conjunto de dados, o primeiro quartil é de 303,5.

5. O terceiro quartil registado, neste conjunto de dados, é de 386.

6. A amplitude interquartil é determinada pela diferença entre o terceiro quartil e o primeiro quartil.
Tem-se que:

$$Amplitude\ interquartil = 386 - 303{,}5 = 82{,}5$$

Este conjunto de dados regista uma amplitude interquartil de 82,5.

Aplicar - Página 120

21. Começamos por ordenar, de forma crescente, o conjunto de pontos do Manuel e o conjunto de pontos do Nuno. Assim:

N.º de pontos registados pelo Manuel

2	4	8	8	9	10	12	12	15	16

N.º de pontos registados pelo Nuno

3	3	4	4	4	4	5	5	5	6

De seguida, recolhemos os valores dos extremos e quartis de cada um dos jogadores da forma seguinte.

N.º de pontos registados pelo Manuel

$mín. = 2$ $máx. = 16$

2	4	8	8	9	10	12	12	15	16
				$M_e = \frac{9+10}{2} = 9{,}5$					

2	4	8	8	9
		$Q_1 = 8$		

10	12	12	15	16
		$Q_3 = 12$		

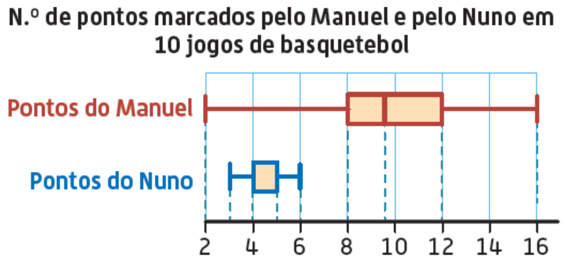
N.º de pontos registados pelo Nuno

3	3	4	4	4	4	5	5	5	6
				$M_e = \frac{4+4}{2} = 4$					

3	3	4	4	4
		$Q_1 = 8$		

4	5	5	5	6
		$Q_3 = 12$		

Procedemos então à construção de dois diagramas de extremos e quartis paralelos para a representação dos dados da forma seguinte.



Fonte: Treinador

Podemos concluir que em, pelo menos, 75 % dos jogos, o Manuel conseguiu marcar 8 pontos ou mais. O Nuno, por sua vez, no máximo conseguiu marcar 6 pontos. Em, pelo menos, 25 % dos jogos, o Nuno marcou 4 ou mais pontos.

O número de pontos marcados pelo Manuel varia entre 2 e 16. Nota-se uma menor variabilidade no número de pontos marcados pelo Nuno, pois os pontos marcados por este jogador concentram-se entre 3 e 6 (inclusive). Já o Manuel apresenta uma maior variabilidade no conjunto de pontos marcados, pois os seus pontos concentram-se no intervalo entre 2 e 16 pontos.

Podemos concluir que o Manuel teve uma melhor prestação nestes dez jogos da temporada.

Aplicar - Página 121

22.

22.1. Tendo em conta o gráfico de barras do número de livros lidos pelos 40 alunos, podemos determinar a média (arredondada às unidades).

$$\bar{x} = \frac{1 \times 10 + 2 \times 8 + 3 \times 6 + 7 \times 4 + 8 \times 3 + 10 \times 1 + 11 \times 3 + 12 \times 5}{40} = \frac{199}{40} = 4{,}975$$

A média do número de livros lidos nas férias de verão, por cada aluno, é, aproximadamente, igual a 5.

Analisando a distribuição dos dados apresentada no gráfico de barras, podemos afirmar que a média não é um bom indicador, pois não houve qualquer aluno a ler 5 livros, 24 alunos leram 3 livros ou menos e 16 alunos leram 7 ou mais livros.

22.2. Como existem 40 alunos em estudo, a mediana irá corresponder à média aritmética do número de livros lidos dos alunos que ocupam as 20.ª e 21.ª posição, quando os dados estão ordenados de forma crescente.

Tem-se:

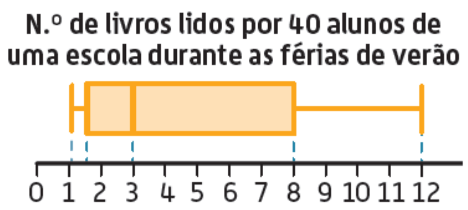
$$M_e = \frac{x_{20} + x_{21}}{2} = \frac{3+3}{2} = 3$$

Os quartis irão corresponder aos 10.º e 11.º alunos dos conjuntos formados pelos 20 primeiros alunos e os 20 últimos alunos, quando estes estão ordenados de forma crescente de livros lidos. Assim:

$$Q_1 = \frac{x_{10}+x_{11}}{2} = \frac{1+2}{2} = 1{,}5 \qquad Q_3 = \frac{x_{30}+x_{31}}{2} = \frac{8+8}{2} = 8$$

Os valores máximos e mínimos de livros lidos pelos alunos são 12 e 1, respetivamente.

Com os extremos e quartis recolhidos, podemos representar os dados no diagrama de extremos e quartis seguinte.



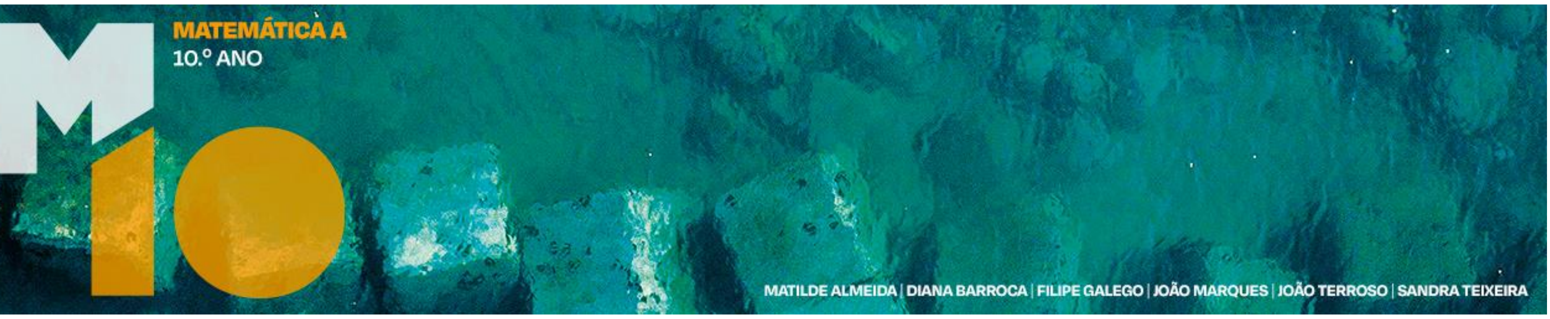
Fonte: Direção da Escola

Esta distribuição de dados é enviesada à direita, pois existe maior concentração de dados entre o mínimo e a mediana.

Tarefa 9 - Página 122

1. Podemos concluir que:

$$Duração\ mínima = 20 \qquad Duração\ máxima = 640$$



A amplitude deste conjunto de dados é:

$$Amplitude = 640 - 20 = 620$$

A amplitude das durações das chamadas feitas pela Olga é 620 segundos.

2. A média da duração das chamadas é determinada por:

$$\bar{x} = \frac{120 + 480 + 320 + 640 + 240 + 20 + 90 + 200 + 140 + 250}{10} = \frac{2500}{10} = 250 \; segundos$$

A duração média, em segundos, das chamadas feitas pela Olga foi de 250 segundos.

3.

3.1. O desvio em relação à média é determinado pelo valor absoluto da diferença entre a duração da chamada e a média do conjunto de durações. Assim, o desvio em relação à média de uma chamada com a duração 90 segundos é:

$$Desvio \; em \; relação \; à \; média = |90 - 250| = |-160| = 160 \; segundos$$

3.2. De modo semelhante à alínea anterior, o desvio em relação à média de uma chamada com a duração 480 segundos é:

$$Desvio \; em \; relação \; à \; média = |480 - 250| = |230| = 230 \; segundos$$

4. Para determinarmos a mediana das durações das chamadas é necessário, em primeiro lugar, ordenar essas durações, por ordem crescente.

20	90	120	140	200	240	250	320	480	640

Como temos um número par de dados, vamos ter dois valores centrais, pelo que:

20	90	120	140	200	240	250	320	480	640

$$M_e = \frac{200 + 240}{2} = 220$$

A duração mediana deste conjunto de dados é 220 segundos.

5. Para procedermos à construção do diagrama de extremos e quartis das durações de chamada registadas pela Olga, deveremos determinar os extremos e quartis deste conjunto de dados.

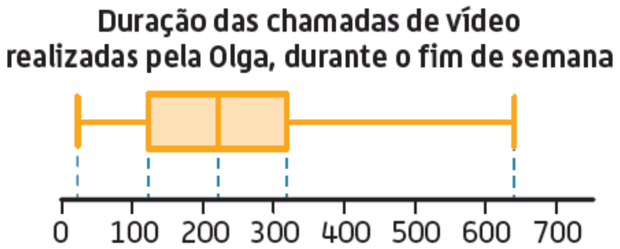
$$M_e = 220$$

$$mín. = 20 \qquad máx. = 640$$

20	90	120	140	200
		$Q_1 = 120$		

240	250	320	480	640
		$Q_3 = 320$		

Com estas medidas, podemos representar o diagrama de extremos e quartis associado da forma seguinte.



Fonte: Própria

Podemos verificar que a distribuição de dados é enviesada à direita, pois existe mais concentração de dados entre a duração mínima e a duração mediana.

A duração mediana, 220 segundos, corresponde ao valor central dos dados.

Nota-se alguma simetria nos 50% dos valores centrais, entre os quartis 1 e 3.

Observa-se alguma dispersão dos dados entre a duração mediana, 220 segundos, e a duração máxima das chamadas, 640 segundos. Existe maior concentração dos dados entre o valor mínimo, 20 segundos de duração, e a duração mediana do conjunto de chamadas, 220 segundos.

Aplicar - Página 123

23.

23.1. Tendo como referência a hora prevista de chegada e a hora de chegada efetiva do conjunto de voos, podemos determinar os atrasos da forma seguinte.

Horas previstas e horas de chegada efetivas de um conjunto de aviões

Hora prevista de chegada	Hora de chegada	Atraso (minutos)
19h00	20h43	$60+43=103$
19h15	21h10	$60+45+10=115$
19h45	20h58	$60+13=73$
19h45	21h18	$60+15+18=93$
20h00	20h37	37
20h00	21h32	$60+32=92$
20h00	22h05	$60+60+5=125$

23.2. A média dos atrasos, em minutos, é: $\bar{x}=\frac{103+115+73+93+37+92+125}{7}\approx 91{,}14$

A média dos atrasos é, aproximadamente, 91,14 minutos.

Aplicar - Página 124

24. Seja y o valor da licitação em falta. Tem-se que:

$$\bar{x} = 34 \Leftrightarrow \frac{14+16+22+31+32+37+45+48+50+y}{10} = 34 \Leftrightarrow$$

$$\Leftrightarrow 14+16+22+31+32+37+45+48+50+y = 340 \Leftrightarrow$$

$$\Leftrightarrow\ 295 + y = 340 \Leftrightarrow\ y = 340 - 295 \Leftrightarrow\ y = 45$$

R.: O valor da licitação em falta é 45 euros.

25. Comecemos por identificar as marcas de classe.

Altura, em centímetros, dos elementos da banda filarmónica de Abraços de Cima

Altura (cm)	Número de elementos	Marca da classe
[130,140[	5	$\frac{130+140}{2} = 135$
[140,150[	7	$\frac{140+150}{2} = 145$
[150,160[	8	$\frac{150+160}{2} = 155$
[160,170[	10	$\frac{160+170}{2} = 165$
[170,180[	15	$\frac{170+180}{2} = 175$
[180,190[	3	$\frac{180+190}{2} = 185$
[190,200[	2	$\frac{190+200}{2} = 195$
Total	50	

Um valor aproximado da média é:

$$\bar{x} = \frac{135\times 5 + 145\times 7 + 155\times 8 + 165\times 10 + 175\times 15 + 185\times 3 + 195\times 2}{50} = 163$$

A média das alturas dos elementos da banda filarmónica pode ser aproximada a 163 centímetros.

Aplicar - Página 125

26.

26.1. O número de golos marcados pelo Benfica varia entre 0 e 4. Podemos construir a tabela de frequências absolutas simples seguinte.

Número de golos marcados pelo Benfica por jogo

N.º de golos	Frequência absoluta
0	1
1	3
2	5
3	2
4	1

26.2. A moda do número de golos marcados neste conjunto de jogos é 2 golos, pois é o dado que regista maior frequência absoluta (5 jogos).

27. A classe modal é a classe [170, 180[, pois é aquela que apresenta maior frequência absoluta.

Aplicar - Página 126

28. Com o conjunto de dados da tabela, podemos concluir que o valor de moda do número de golos é 2, pois corresponde ao dado com maior frequência absoluta registada.
No que diz respeito à mediana, esta irá assumir o valor numérico do 5.º dado, quando o conjunto de golos sofridos é registado por ordem crescente. Assim:

Valor central

0	2	2	2	2	2	4	5	5

O valor da mediana é: $M_e = 2$
O valor da média é: $\bar{x} = \frac{0\times1+2\times5+4\times1+5\times1}{9} \approx 2{,}7$
Concluímos que, no conjunto de golos sofridos, $M_o = 2$, $M_e = 2$ e $\bar{x} \approx 2{,}7$.

29. Começamos por ordenar os dados do gráfico da forma seguinte.

21 000	21 400	a	a	22 200	22 400	22 600	22 600

Como o número de dados é par, a mediana é a média dos dois valores centrais desta lista ordenada. Neste caso, a e 22 200. Podemos, então, calcular o valor de a.

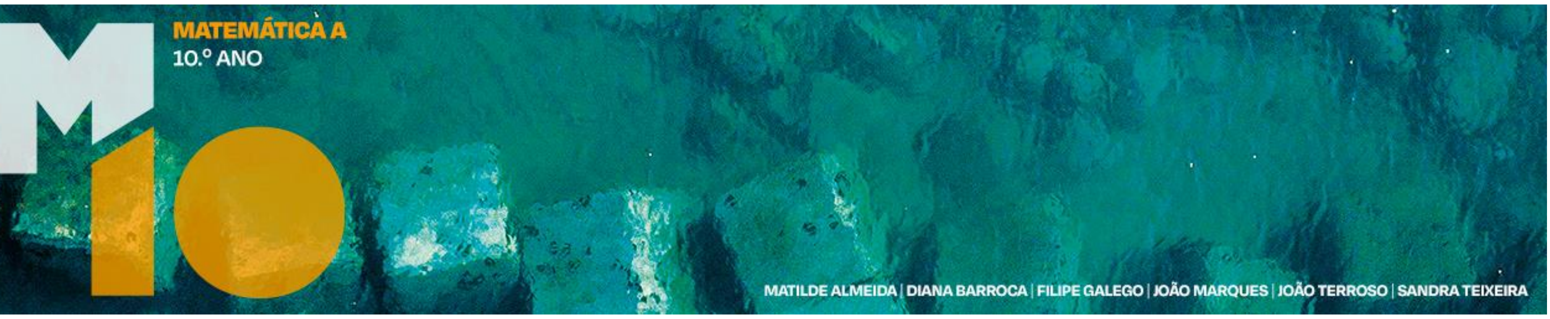
$$\frac{a + 22\,200}{2} = 22\,000 \Leftrightarrow a + 22\,200 = 44\,000 \Leftrightarrow a = 44\,000 - 22\,200 \Leftrightarrow a = 21\,800$$

O valor de a é 21 800.

Aplicar - Página 128
30.
30.1.

Número de iogurtes retirados diariamente

Classes	Frequência absoluta	Frequência absoluta acumulada	Frequência relativa (%)	Frequência relativa acumulada (%)
[0,20[	7	7	$\frac{7}{23} \times 100 \approx 23$	23
[20,40[	9	16	$\frac{9}{30} \times 100 = 30$	$23 + 30 = 53$
[40,60[	3	19	$\frac{3}{30} \times 100 = 10$	$53 + 10 = 63$
[60,80[	5	24	$\frac{5}{30} \times 100 \approx 17$	$63 + 17 = 80$
[80,100[	6	30	$\frac{6}{30} \times 100 = 20$	$80 + 20 = 100$
Total	30		100	



30.2. A mediana corresponde à média aritmética dos dois dados que ocupam as 15.ª e 16.ª posições. Esses dados estão inseridos na classe $[20,40[$, a classe mediana.
A classe modal é $[20,40[$, pois corresponde à classe que regista a maior frequência absoluta.
Concluímos então que a classe modal e a classe mediana coincidem, sendo ambas [20,40[.

Aplicar - Página 129

31.

31.1. A média do conjunto de questões de aula do Manuel é: $\bar{x} = \frac{18+19+20+19+10}{5} = \frac{86}{5} = 17{,}2$
Para determinar a mediana, teremos que ordenar os dados por ordem crescente e encontrar o dado que ocupa a posição central (neste caso, a 3.ª posição).

Assim:

10	18	19	19	20
		$M_e = 19$		

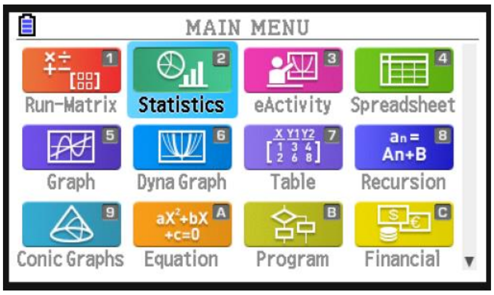
Concluímos então que a média das questões de aula do Manuel é 17,2 valores e a mediana é 19 valores.

31.2. Nesta situação, a mediana parece ser a mais adequada, porque é menos influenciada por um valor (o da questão de aula 5) em que o Manuel obteve uma classificação muito diferente das suas classificações habituais. Essa classificação, por ser muito díspar das restantes, causa um grande impacto na média do Manuel.

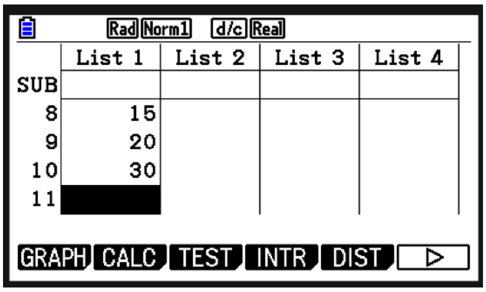
Aplicar - Página 131

32.

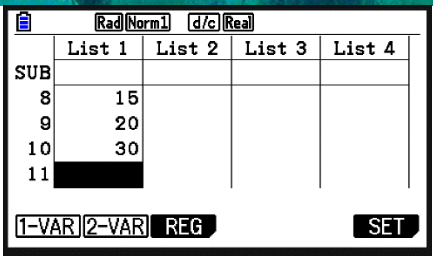
Recorrendo às potencialidades da calculadora gráfica (Casio CG-50), começamos por, no menu inicial, selecionar a opção Estatística.



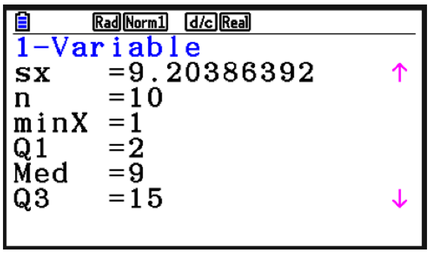
De seguida, introduzimos os dados na lista 1.



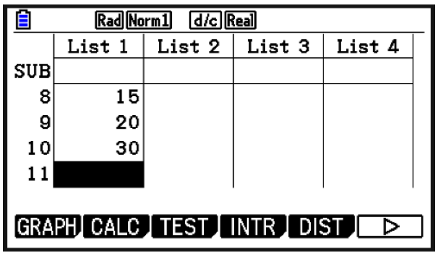
Depois, selecionamos F2 (CALC) e verificamos as definições para o cálculo das medidas estatísticas F6 (SET). Nas definições, colocamos 1Var XList: List 1 e em 1Var Freq: 1 (queremos admitir que a frequência absoluta de cada um dos dados da lista 1 é de 1 unidade).

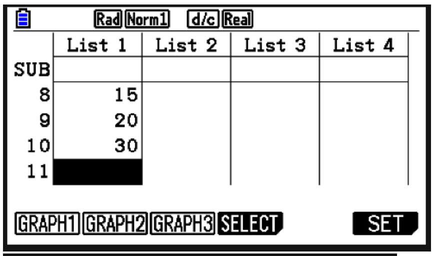


Posteriormente, selecionamos EXIT e carregamos em F1 (1-VAR). Obtemos então a janela seguinte.

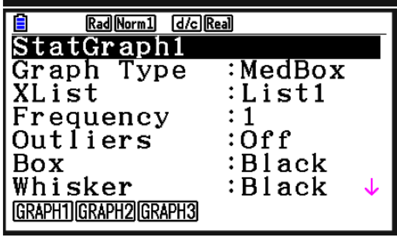


Para obter o diagrama de extremos e quartis (no mesmo modelo de calculadora), percorrem-se os passos seguintes.
Com a mesma lista 1, selecionamos F1 (GRAPH).

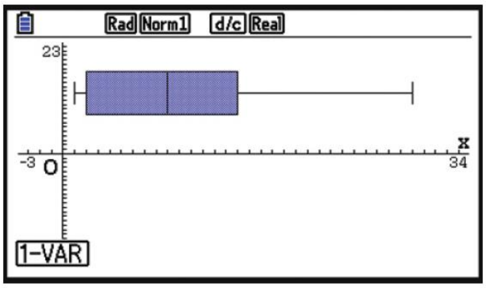




E depois selecionamos F6 (SET). No tipo de gráfico, optamos por *MedBox* (diagrama de extremos e quartis), definimos List1 na *XList* e 1 na *Frequency*.

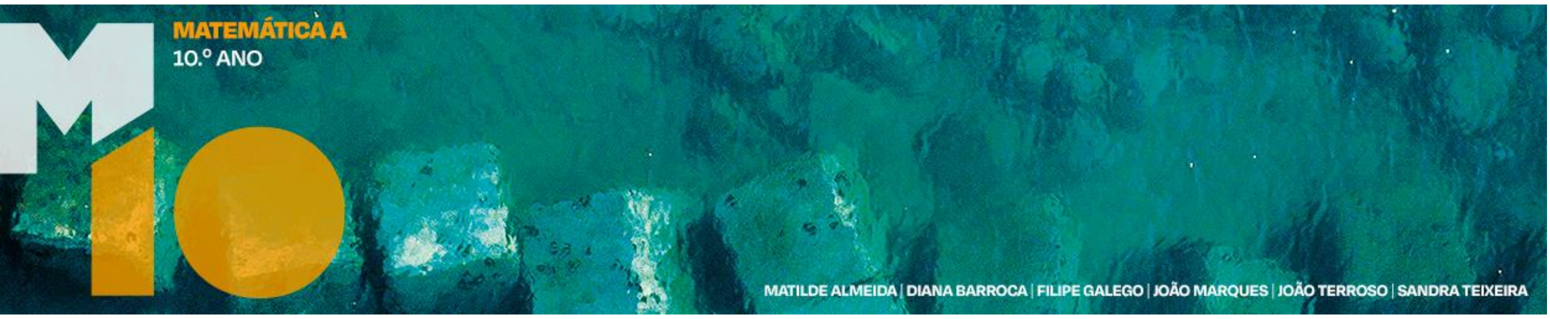


De seguida, carregamos em GRAPH1, obtendo o diagrama de extremos e quartis seguinte.



$Q_1 = 2$; $M_e = 9$; $Q_3 = 15$

Nota: Poderemos ter de configurar a janela de visualização da máquina. Para isso, selecionamos SHIFT, seguido de F3 (V-Window).



Aplicar - Página 132

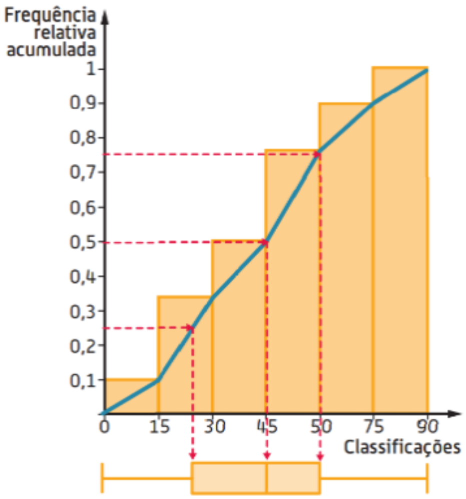
33. Para determinar graficamente a mediana e os quartis associados a este conjunto de classificações, temos de recorrer à frequência relativa acumulada de cada uma das classes.

Classificações obtidas na Prova Final de Matemática, do 3.º Ciclo, pelos alunos do 9.º ano de uma escola

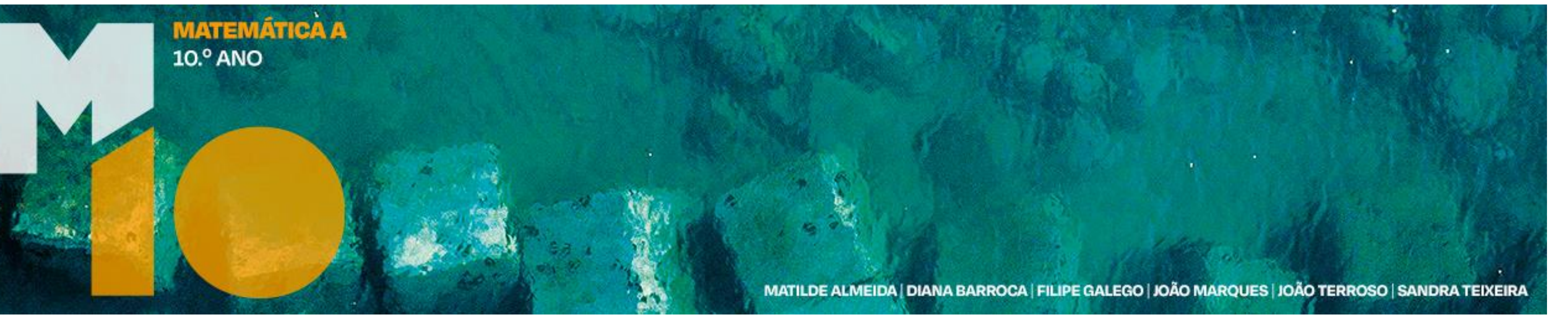
Classificações	Número de alunos	Frequência relativa (%)	Frequência relativa acumulada (%)
[0,15[	5	$\frac{5}{50} \times 100 = 10$	10
[15,30[	12	$\frac{12}{50} \times 100 = 24$	$24 + 10 = 34$
[30,45[	8	$\frac{8}{50} \times 100 = 16$	$34 + 16 = 50$
[45,60[	13	$\frac{13}{50} \times 100 = 26$	$50 + 26 = 76$
[60,75[	7	$\frac{7}{50} \times 100 = 14$	$76 + 14 = 90$
[75,90[	5	$\frac{5}{50} \times 100 = 10$	$90 + 10 = 100$
Total	50	100	

De seguida, e no mesmo referencial, representamos o histograma de frequências relativas acumuladas, o polígono de frequências relativas acumuladas e o diagrama de extremos e quartis seguintes.

Histograma e diagrama de extremos e quartis das classificações obtidas na Prova Final de Matemática, do 3.º Ciclo, dos alunos de uma escola



Fonte: Própria



Aplicar - Página 133

34.

34.1. Traçando uma linha auxiliar perpendicular ao eixo horizontal a partir de 11 kg e intersetando com o polígono de frequências relativas acumuladas, podemos concluir que o peso de 11 kg corresponde ao percentil 50, P_{50}.

34.2. Podemos afirmar que 50 % das meninas com 22 meses têm massa menor ou igual a 11 kg.

34.3. O percentil 50, P_{50}, pode também ser designado por mediana.

Aplicar - Página 134

35.

35.1. Para facilitar a ordenação dos dados, podemos recorrer a um diagrama de caule-e-folhas. Assim:

Tempo, em minutos, de frequência da biblioteca escolar

0	2 5 6 7 8 8
1	0 0 1 8 8 9
2	0 2 2 4 5
3	0 4 5 7 9 9
4	0 0

0|2 significa 2 minutos

Fonte: Registos da Mariana

Podemos dispor os dados por ordem crescente da forma seguinte.

2	5	6	7	8	8	10	10	11	18	18	19	20	22	22	24	25	30	34	35	37	39	39	40	40

Os percentis 25, 50 e 75 correspondem aos valores do primeiro quartil, da mediana e do terceiro quartil. Assim:

2	5	6	7	8	8	10	10	11	18	18	19	20	22	22	24	25	30	34	35	37	39	39	40	40
					Q_1							M_e						Q_3						

$$Q_1 = \frac{8+10}{2} = 9 \qquad Q_3 = \frac{34+35}{2} = 34{,}5$$

Concluímos, então, que $P_{25} = Q_1 = 9$, $P_{50} = M_e = 20$ e $P_{75} = Q_3 = 34{,}5$.

35.2. Iremos recorrer à tabela de frequências relativas acumuladas associada a este conjunto de dados para determinar o menor valor dos 30% dos registos de maior valor.

Tempo de frequência, em minutos, na biblioteca escolar

Tempo (em minutos)	Frequência absoluta	Frequência relativa (%)	Frequência relativa acumulada (%)
2	1	$\frac{1}{25}\times 100=4$	4
5	1	$\frac{1}{25}\times 100=4$	$4+4=8$
6	1	$\frac{1}{25}\times 100=4$	$8+4=12$
7	1	$\frac{1}{25}\times 100=4$	$12+4=16$
8	2	$\frac{2}{25}\times 100=8$	$16+8=24$
10	2	$\frac{2}{25}\times 100=8$	$24+8=32$
11	1	$\frac{1}{25}\times 100=4$	$32+4=36$
18	2	$\frac{2}{25}\times 100=8$	$36+8=44$
19	1	$\frac{1}{25}\times 100=4$	$44+4=48$
20	1	$\frac{1}{25}\times 100=4$	$48+4=52$
22	2	$\frac{2}{25}\times 100=8$	$52+8=60$
24	1	$\frac{1}{25}\times 100=4$	$60+4=64$
25	1	$\frac{1}{25}\times 100=4$	$64+4=68$
30	1	$\frac{1}{25}\times 100=4$	$68+4=72$
34	1	$\frac{1}{25}\times 100=4$	$72+4=76$
35	1	$\frac{1}{25}\times 100=4$	$76+4=80$
37	1	$\frac{1}{25}\times 100=4$	$80+4=84$
39	2	$\frac{2}{25}\times 100=8$	$84+8=92$
40	2	$\frac{2}{25}\times 100=8$	$92+8=100$
Total	**25**	**100**	

Como $100\%-30\%=70\%$, temos que o menor valor de entre os 30% de maiores tempos registados é o P_{70}. Consultando a tabela de frequências, $P_{70}=30$.
Do conjunto dos 30% maiores tempos registados pela Mariana, o que teve menor duração foi 30 minutos.

35.3. Existem dois tempos de 40 minutos registados pela Mariana.
Ou seja:

$$Percentagem = \frac{2}{25} \times 100 = 8\%$$

O tempo 40 minutos pertence a $P_{93}, P_{94}, P_{95}, P_{96}, P_{97}, P_{98}$ e P_{99}.

Aplicar - Página 135

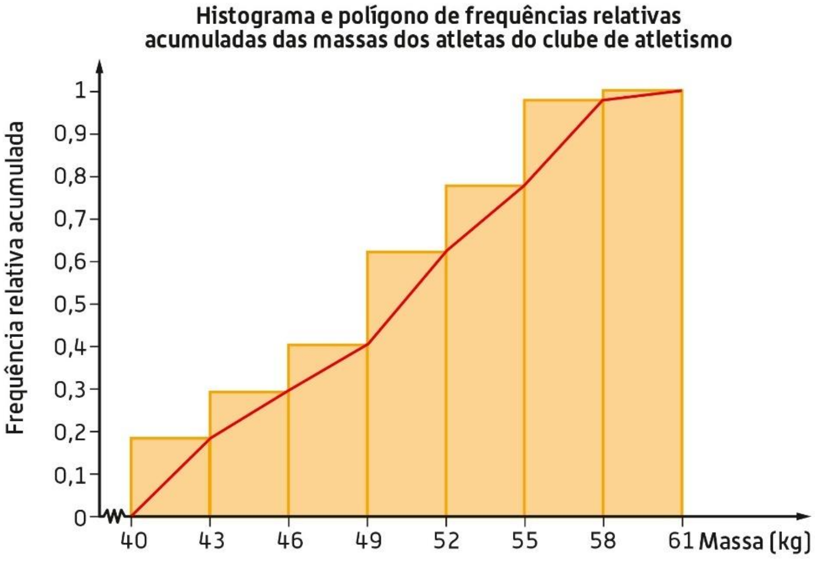
36.

36.1. Começamos por elaborar uma tabela de frequências relativas acumuladas semelhante à seguinte.

Massa, em quilogramas, dos atletas de um clube de atletismo

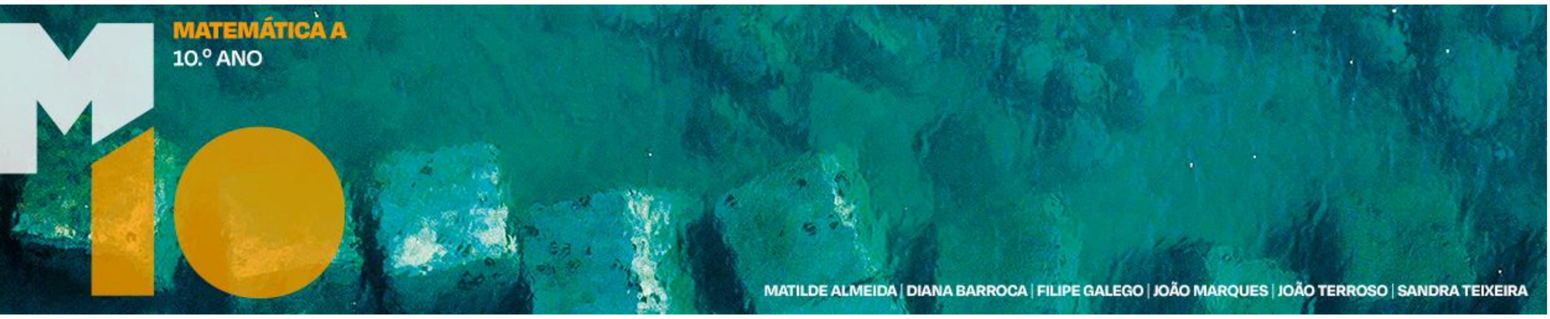
Massa (kg)	Número de atletas	Frequência relativa (%)	Frequência relativa acumulada (%)
$[40, 43[$	9	18	18
$[43, 46[$	6	12	30
$[46, 49[$	7	14	44
$[49, 52[$	9	18	62
$[52, 55[$	8	16	78
$[55, 58[$	7	14	92
$[58, 61[$	4	8	100
Total	50	100%	

Depois, elaboramos o histograma e o polígono de frequências relativas acumuladas.



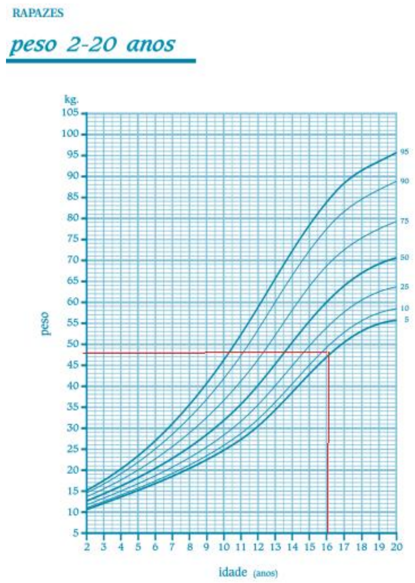
Fonte: Própria

A massa do atleta Manuel está no P_{40}, aproximadamente.



36.2.

Consultando os percentis da massa, em quilogramas, da população portuguesa até aos 20 anos, indicados pela Direção Geral de Saúde, a massa do atleta Manuel está no percentil P_5.
O valor do percentil da massa do atleta Manuel é P_5 dentro dos valores registados na população portuguesa. Já de entre os valores da equipa de atletismo, o Manuel atinge o P_{40}. Assim, o valor do percentil do Manuel dentro da população portuguesa é menor quando comparado com o percentil deste atleta dentro da equipa.



Fonte: Direção Geral de Saúde

Tarefa 10 - Página 136

1. Usando os registos efetuados pelo Frederico nos seus treinos diários, podemos construir o diagrama de caule-e-folhas seguinte.

Número de piscinas realizadas em cada dia pelo Frederico

2	0 0 5
3	0 0 0 0 5
4	0 0 0 0 0 5
5	0 0 0 0 0
6	0

2|0 significa 20

Fonte: Registos do Frederico

2. Com o diagrama de caule-e-folhas, podemos ordenar, de forma crescente, os dados da forma seguinte.

20	20	25	30	30	30	30	35	40	40	40	40	40	45	50	50	50	50	50	60

Podemos então determinar os extremos e quartis do conjunto de dados registados pelo Frederico durante os seus treinos diários.

$$mín. = 20 \qquad máx. = 60$$

20	20	25	30	30	30	30	35	40	40	40	40	40	45	50	50	50	50	50	60

$$M_e = \frac{40+40}{2} = 40$$

20	20	25	30	30	30	30	35	40	40

$$Q_1 = \frac{30 + 30}{2} = 30$$

40	40	40	45	50	50	50	50	50	60

$$Q_3 = \frac{50 + 50}{2} = 50$$

As amplitude e amplitude interquartil deste conjunto de dados são:

- $Amplitude = 60 - 20 = 40$
- $Amplitude\ Interquartil = Q_3 - Q_1 = 50 - 30 = 20$

O valor da amplitude é igual a 40 e o valor da amplitude interquartil é igual a 20.

3. O número de piscinas médio que o Frederico registou nestes 20 dias de treino é:

$$\bar{x} = \frac{2 \times 20 + 25 + 4 \times 30 + 35 + 5 \times 40 + 45 + 5 \times 50 + 60}{20} = \frac{775}{20} = 38{,}75$$

O número médio de piscinas realizadas pelo Frederico por dia foi de 38,75.

4. A média deste conjunto de dados é $\bar{x} = 38{,}75$. Seguindo as indicações, podemos então completar a tabela da forma seguinte.

Número de piscinas (x_i)	Frequência absoluta (n_i)	Desvio para a média $d_i = x_i - \bar{x}$	$d_i^{\,2}$	$n_i \times d_i^{\,2}$
20	2	$20 - 38{,}75 = -18{,}75$	351,5625	$2 \times 351{,}5625 = 703{,}125$
25	1	$25 - 38{,}75 = -13{,}75$	189,0625	$1 \times 189{,}0625 = 189{,}0625$
30	4	$30 - 38{,}75 = -8{,}75$	76,5625	$4 \times 76{,}5625 = 306{,}25$
35	1	$35 - 38{,}75 = -3{,}75$	14,0625	$1 \times 14{,}0625 = 14{,}0625$
40	5	$40 - 38{,}75 = 1{,}25$	1,5625	$5 \times 1{,}5625 = 7{,}8125$
45	1	$45 - 38{,}75 = 6{,}25$	39,0625	$1 \times 39{,}0625 = 39{,}0625$
50	5	$50 - 38{,}75 = 11{,}25$	126,5625	$5 \times 126{,}5625 = 632{,}8125$
60	1	$60 - 38{,}75 = 21{,}25$	451,5625	$1 \times 451{,}5625 = 451{,}5625$
Total	**20**			**2343,75**

5.

5.1. Temos que:

$s^2 = \frac{2343{,}75}{20-1} \approx 123{,}35$ A variância é, aproximadamente, 123,35.

5.2. O desvio padrão deste conjunto de dados é:

$s = \sqrt{123{,}35} \approx 11{,}11$ O desvio padrão é, aproximadamente, 11,11.

Tarefa 11 – Página 136

O programa apresentado permite apenas determinar os valores máximo (identificado com a variável *máximo* nas linhas de código) e mínimo (identificado com a variável *mínimo* nas linhas de código). O utilizador é apenas informado sobre os dados estatísticos máximo e mínimo.

```
# programa para determinar a amplitude
import math
idades=[14,16,14,15,17,23]
idades.sort()
maximo=idades[5]
minimo=idades[0]
print("O mínimo é ",minimo,".")
print("O máximo é ",maximo,".")
```

Para que o utilizador seja informado da amplitude associada a este conjunto de dados, teremos de, por exemplo, adicionar um novo *print* que faça o cálculo desse valor. Sugere-se o programa seguinte como complemento ao que é apresentado.

```
# programa para determinar a amplitude
import math
idades=[14,16,14,15,17,23]
idades.sort()
maximo=idades[5]
minimo=idades[0]
amplitude=maximo-minimo
print("O mínimo é ",minimo,".")
print("O máximo é ",maximo,".")
print("A amplitude deste conjunto de dados é ",amplitude,".")
```

Fazendo um teste ao programa escrito, obtém-se o seguinte.

```
# programa para determinar a amplitude
import math
idades=[14,16,14,15,17,23]
idades.sort()
maximo=idades[5]
minimo=idades[0]
amplitude=maximo-minimo
print("O mínimo é ",minimo,".")
print("O máximo é ",maximo,".")
print("A amplitude deste conjunto de dados é ",amplitude,".")
```

```
O mínimo é  14 .
O máximo é  23 .
A amplitude deste conjunto de dados é  9 .
```

Aplicar - Página 137

37.

37.1. A variável em estudo é o número de medalhas conquistadas por Portugal nas várias edições dos Jogos Olímpicos. É uma variável quantitativa discreta.

37.2. O número de medalhas recebidas pelo comité português nas 8 edições de jogos olímpicos registados variou entre 0 e 4.
Com os dados registados, conseguimos construir a seguinte tabela de frequências simples.

Número de medalhas conquistadas por Portugal em 8 edições de Jogos Olímpicos

Número de medalhas	Frequência absoluta	Frequência relativa
0	1	$\frac{1}{8} = 0{,}125$
1	2	$\frac{2}{8} = 0{,}25$
2	3	$\frac{3}{8} = 0{,}375$
3	1	$\frac{1}{8} = 0{,}125$
4	1	$\frac{1}{8} = 0{,}125$
Total	8	1

37.3. Representando o conjunto de dados, por ordem crescente, obtemos o seguinte.

0	1	1	2	2	2	3	3

Os extremos e quartis são:

$Valor\ mínimo = 0$ $\qquad$ $Valor\ máximo = 4$

0	1	1	2	2	2	3	3

$$M_e = \frac{2+2}{2} = 2$$

0	1	1	2

$$Q_1 = \frac{1+1}{2} = 1$$

2	2	3	3

$$Q_3 = \frac{2+3}{2} = 2{,}5$$

Podemos determinar os valores da amplitude e da amplitude interquartil:

- $Amplitude = Valor\ máximo - Valor\ mínimo = 4 - 0 = 4$
- $Amplitude\ interquartil\ = Q_3 - Q_1 = 2{,}5 - 1 = 1{,}5$

A amplitude deste conjunto de dados é 4 e a amplitude interquartil é 1,5.

Aplicar - Página 138

38.

38.1. O tempo médio de minutos jogados por João Mário é:

$$\bar{x} = \frac{46 + 92 + 82 + 71 + 72 + 103 + 100}{7} \approx 80{,}86$$

Concluímos que o tempo médio de jogo de João Mário foi de, aproximadamente, 80,86 minutos.

38.2. A soma dos quadrados dos desvios em relação à média é:

$Soma\ dos\ quadrados\ dos\ desvios =$

$= (46 - 80{,}86)^2 + (92 - 80{,}86)^2 + (82 - 80{,}86)^2 + \cdots + (103 - 80{,}86)^2 + (100 - 80{,}86)^2 \approx 2372{,}86$

A variância é obtida por:

$$s^2 = \frac{2372{,}86}{7 - 1} \approx 395{,}48$$

O valor da variância associada a este conjunto de dados é, aproximadamente, 395,48.

38.3. O desvio padrão corresponde a:

$s = \sqrt{395{,}48} \approx 19{,}89$

O desvio padrão é, aproximadamente, 19,89.

Aplicar - Página 139

39. O valor médio do IMC na turma C é:

$$\bar{x} = \frac{15 + 16 + 17 + 18 \times 3 + 19 \times 2 + 20 \times 3 + 21 \times 3 + 25 + 28 + 29 + 30 + 31 \times 2 + 32 + 34}{22} \approx 22{,}9$$

A soma dos quadrados dos desvios é:

$Soma = (15 - 22{,}9)^2 + (16 - 22{,}9)^2 + (17 - 22{,}9)^2 + \cdots + 2 \times (31 - 22{,}9)^2 + (32 - 22{,}9)^2 + (34 - 22{,}9)^2 \approx 738{,}6$

A variância e o desvio padrão são determinados por:

- $s^2 = \frac{738{,}6}{22-1} \approx 35{,}2$
- $s = \sqrt{35{,}2} \approx 5{,}9$

Concluímos que $\bar{x} \approx 22{,}9$, $s^2 \approx 35{,}2$ e $s \approx 5{,}9$.

40. Começamos por determinar a marca da classe de cada uma das classes consideradas.

Distância (em km)	Frequência absoluta	Marca da classe
[0,6[	7	$\frac{0+6}{2}=3$
[6,12[	6	$\frac{6+12}{2}=9$
[12,18[	5	$\frac{12+18}{2}=15$
[18,24[	2	$\frac{18+24}{2}=21$
[24,30[	1	$\frac{24+30}{2}=27$
Total	21	

Um valor aproximado para a média destas distâncias é:

$$\bar{x}=\frac{3\times 7+9\times 6+15\times 5+21\times 2+27\times 1}{21}\approx 10{,}429$$

Temos que um valor aproximado da soma dos quadrados dos desvios é:

$Soma =$

$$=7\times(3-10{,}429)^2+6\times(9-10{,}429)^2+5\times(15-10{,}429)^2+2\times(21-10{,}429)^2+1\times(27-10{,}429)^2\approx 1001{,}143$$

Concluímos então que a variância e o desvio padrão são:

- $s^2=\frac{1001{,}143}{21-1}\approx 50{,}057$
- $s=\sqrt{50{,}057}\approx 7{,}08$

O valor do desvio padrão é, aproximadamente 7,08 quilómetros.

Tarefa 12

1.

Reproduzindo as células da coluna A, entre as linhas 2 e 7, e triplicando esses valores na coluna B, obtemos a tabela seguinte.

	A	B
1	a	b
2	2	6
3	3	9
4	4	12
5	5	15
6	6	18
7	7	21

De seguida, usamos a ferramenta MEDIA do Excel para determinar a média de B.

Assim:

	A	B
1	a	b
2	2	6
3	3	9
4	4	12
5	5	15
6	6	18
7	7	21
8		
9	4,5	13,5

Comparando os valores das médias obtidas, podemos concluir que o valor da média obtido em B9 (4,5) é igual ao triplo do valor da média obtido em A9 ($3 \times 4{,}5 = 13{,}5$).

2. Procedendo de forma semelhante, mas multiplicando cada um dos dados estatísticos por 5, obtemos os seguintes dados e médias na coluna B:

	A	B
1	a	b
2	2	10
3	3	15
4	4	20
5	5	25
6	6	30
7	7	35
8		
9	4,5	22,5

Conseguimos então observar que o valor da média obtido em B9 (4,5) é igual ao quíntuplo do valor da média obtido em A9 ($5 \times 4{,}5 = 22{,}5$).

Tarefa 13

1. Resposta ao cuidado do aluno.
2. Reproduzindo no *Excel* as classificações registadas nos primeiros e segundos testes, obtemos a tabela seguinte.

	A	B	C	D	E
1		Classificações dos primeiros testes		Classificações dos segundos testes	
2		14		16	
3		15		17	
4		16		18	
5		17		19	

De seguida, recorremos à ferramenta MEDIA do Excel, para determinar as médias de cada um dos conjuntos de testes:

	A	B	C	D
1		Classificações dos primeiros testes		Classificações dos segundos testes
2		14		16
3		15		17
4		16		18
5		17		19
6				
7	Média	=MÉDIA(B2:B5)		=MÉDIA(D2:D5)

	A	B	C	D
1		Classificações dos primeiros testes		Classificações dos segundos testes
2		14		16
3		15		17
4		16		18
5		17		19
6				
7	Média	15,5		17,5

A média dos primeiros testes é de 15,5 valores e a média dos segundos testes é de 17,5 valores. A média do João entre os primeiros testes e os segundos testes aumentou 2 valores.

Aplicar - Página 141

41.

41.1. A média do conjunto dos 10 goleadores da liga portuguesa de hóquei patins nessa época é:

$$\bar{x} = \frac{53 + 45 + 43 + 38 + 37 + 36 \times 2 + 31 + 28 + 27}{10} = 37{,}4$$

A média do número de golos marcados pelos 10 jogadores é 37,4 golos.

41.2. Caso o campeonato tivesse mais um jogo e nele cada um destes goleadores marcasse dois golos, (o que corresponderia ao aumento de duas unidades a cada um dos dados em estudo), a média do número de golos marcados seria igual à média anterior adicionada de duas unidades.

Ou seja:

$$\bar{y} = \bar{x} + 2 = 37{,}4 + 2 = 39{,}4$$

A média do número de golos marcados pelos 10 jogadores passaria a ser 39,4 golos.

Aplicar - Página 142

42.

42.1. A média das classificações na disciplina de matemática foi:

$$\bar{x} = \frac{9 + 12 + 13 + 16 + 18 + 19 + 20 + 20 + \cdots + 12 + 15 + 12}{19} = 14$$

A soma dos quadrados dos desvios em relação à média é:

$$soma = (9 - 14)^2 + (12 - 14)^2 + (13 - 14)^2 + \cdots + (15 - 14)^2 + (12 - 14)^2 = 346$$

- A variância do conjunto de classificações é:
  - $s^2 = \frac{346}{19-1} \approx 19{,}222$
- O valor do desvio padrão é:
  - $s = \sqrt{19{,}222} \approx 4{,}38$

O desvio padrão das classificações foi, aproximadamente, 4,38 valores.

42.2. No segundo momento de avaliação, as classificações registadas pelos alunos da turma do Bernardo foram:

Classificações (segundo momento)	7	17	13	12	10	18	6	10	11	18	6	13	14	16	5	10	16	18	8

A média da turma também desceu duas unidades, sendo que $\bar{y} = \bar{x} - 2 = 12$.
A soma dos desvios dos quadrados não se alterou:

$$soma = (7-12)^2 + (10-12)^2 + (11-12)^2 + \cdots + (13-12)^2 + (10-12)^2 = 346$$

Logo, a variância e o desvio padrão mantiveram o seu valor, ou seja:

- $s^2 \approx 19{,}222$
- $s \approx 4{,}38$

O desvio padrão das classificações de matemática no segundo momento de avaliação foi, aproximadamente, 4,38 valores.

42.3. As classificações do terceiro momento de avaliação de matemática foram:

Classificações (terceiro momento)	9	19	15	14	12	20	8	12	13	20	8	15	16	18	1	12	18	20	10

No terceiro momento de avaliação, a média das classificações foi:

$$\bar{x} = \frac{1+9+12+13+16+18+19+20+20+\cdots+12+15+12}{19} \approx 13{,}68$$

A soma dos quadrados dos desvios em relação à média é:

$$soma = (1-13{,}68)^2 + (9-13{,}68)^2 + (12-14)^2 + \cdots + (15-14)^2 + (12-14)^2 \approx 464{,}11$$

- A variância do conjunto de classificações é:
  - $s^2 = \frac{464{,}11}{19-1} \approx 25{,}784$
- O valor do desvio padrão é:
  - $s = \sqrt{25{,}784} \approx 5{,}08$

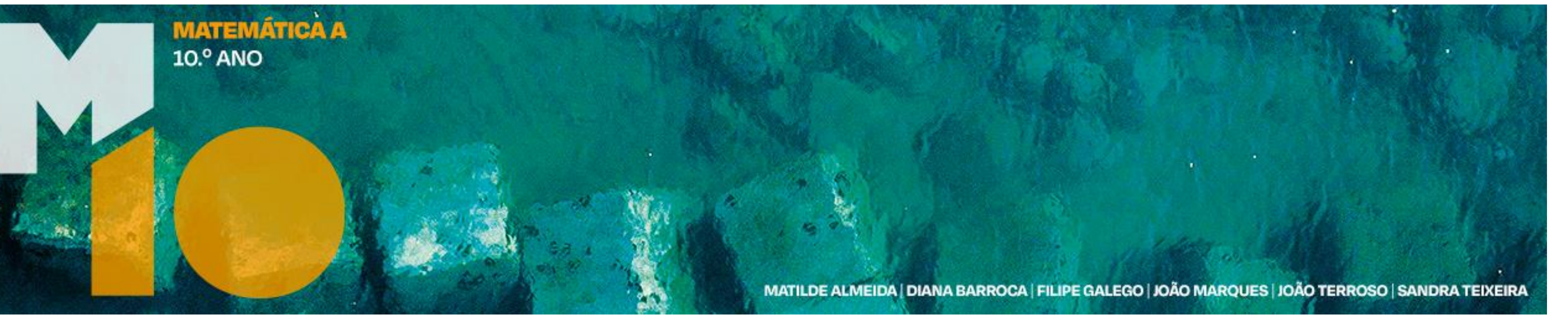
A classificação 1 é muito pequena relativamente aos resultados globais da turma.

Com o registo desse novo valor, a média das classificações diminuiu e o desvio padrão aumentou.

Aplicar – Página 143

43.
Média: $\bar{x} = \frac{3+1+8+3+5}{5} = \frac{20}{5} = 4$

$$\text{Soma dos desvios em relação à média } = (3-4)+(1-4)+(8-4)+(3-4)+(5-4) =$$
$$= -1-3+4-1+1 = 0$$



Tarefa 14 – Página 144

1. A média não representa bem as classificações do José, pois, com exceção da última, essas classificações são superiores à média.
O valor da mediana, 19, representa melhor estes dados, pois sendo o valor central dos dados não foi afetado pelo mau resultado do último teste de mandarim do José.

2. Começamos por ordenar, de forma crescente, as classificações dos testes de mandarim do José, conseguimos determinar os extremos e quartis associados a estes dados.

2	18	19	19	20	20	20
	$Q_1 = 18$		$M_e = 19$		$Q_3 = 20$	

Temos então que:

$$s \approx 6{,}6 \qquad amplitude\ interquartil = Q_3 - Q_1 = 20 - 18 = 2$$

Tendo em conta que o José, numa das avaliações, obteve uma classificação muito inferior às restantes, a amplitude interquartil não representa a variabilidade dos dados. Em contrapartida, o desvio padrão das classificações representa bem melhor essa variabilidade dos dados.

Aplicar – Página 144

44.

44.1. A média do número de reparações efetuadas pelo Eduardo foi:

$$\bar{x} = \frac{1 \times 1 + 2 \times 6 + 3 \times 3 + 4 \times 2 + 18 \times 1 + 22 \times 1}{14} = 5$$

Para determinar a mediana, começamos por escrever os dados, em ordem crescente.

1	2	2	2	2	2	2	3	3	3	4	4	18	22

A mediana corresponde à média aritmética dos números de reparações que ocupam as posições sete e oito, na ordenação.

1	2	2	2	2	2	2	3	3	3	4	4	18	22
						$M_e = \frac{2+3}{2} = 2{,}5$							

Podemos então concluir que a mediana é a medida de localização mais adequada. Tem-se que 12 reparações das 14 registadas pelo Eduardo, são inferiores ao número médio de reparações.

44.2. Tem-se que a soma dos desvios dos quadrados em relação à média deste conjunto de dados é:

$$Soma = (1-5)^2 + 6 \times (2-5)^2 + 3 \times (3-5)^2 + 2 \times (4-5)^2 + (18-5)^2 + (22-5)^2 = 542$$

O desvio padrão é determinado do modo seguinte.

- $s^2 = \frac{542}{14-1} \approx 41{,}692$
- $s = \sqrt{41{,}692} \approx 6{,}45$

Para determinar os quartis, começamos por escrever os dados, em ordem crescente.

1	2	2	2	2	2	2	3	3	3	4	4	18	22

- $Q_1 = 2$ e $Q_3 = 4$ e
- $Amplitude\ interquartil = Q_3 - Q_1 = 4 - 2 = 2$

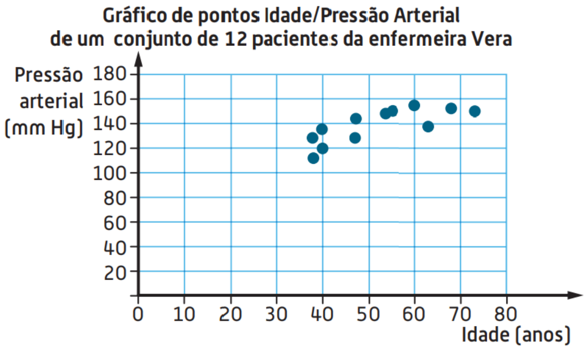
A medida de dispersão mais adequada para descrever este conjunto de dados é a amplitude interquartil. Os dois dias em que o José atendeu a 18 e a 22 ocorrências afetaram muito o valor do desvio padrão.

2.3. Dados bivariados

Tarefa 15 - Página 145

1.

Considerando pares ordenados (x_i, y_i), em que x_i corresponde à variável idade (em anos) e y_i corresponde à variável pressão arterial (em mmHg), obtemos a representação gráfica seguinte.



Fonte: Própria

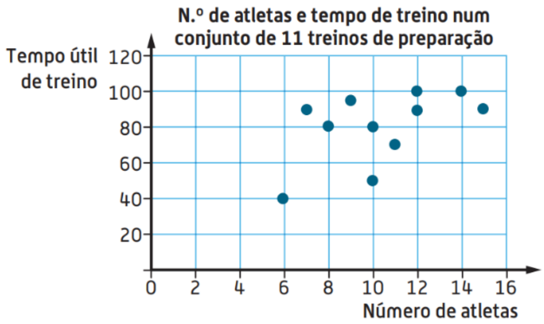
2. Ao observarmos o gráfico, parece haver uma associação entre as variáveis idade e pressão arterial. Os pacientes mais velhos têm tendência a ter uma maior pressão arterial.

3. A variável independente é a idade. A variável dependente é a pressão arterial.

Aplicar – Página 145

45.

45.1. Considerando os pares ordenados (x_i, y_i), em que x_i corresponde à variável número de atletas e y_i corresponde ao tempo útil de treino (em minutos), representamos o conjunto de pontos da forma seguinte.



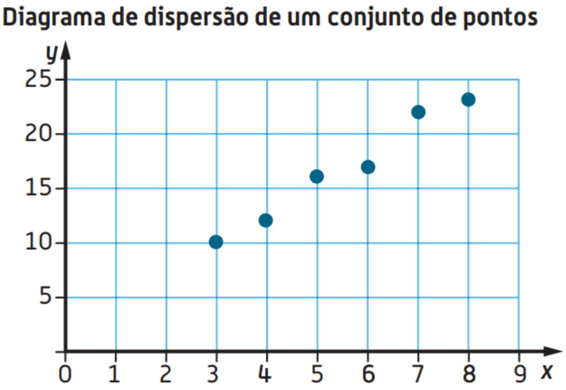
Fonte: Própria

45.2. O número de atletas corresponde à variável independente e o tempo útil de treino é a variável dependente.

45.3. Ao observar o conjunto de pontos representado anteriormente, notamos que, aparentemente, o tempo útil de treino parece ter tendência a ser maior quando o número de atletas é maior.

Aplicar - Página 146

46. Considerando os pares ordenados (x_i, y_i), representamos o diagrama de dispersão da forma seguinte.

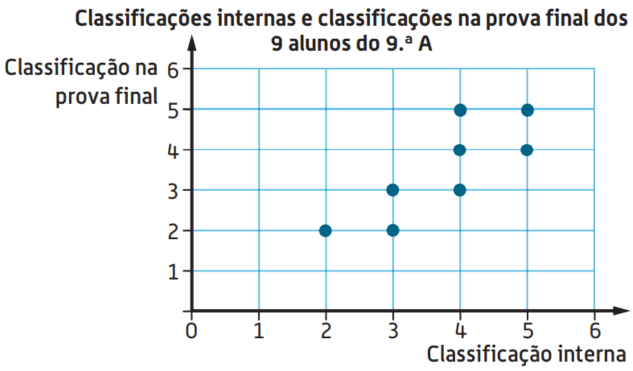


Fonte: Própria

Aplicar - Página 147

47.

47.1. Seja (x_i, y_i) o conjunto de pares ordenados em que x_i corresponde à variável classificação interna e y_i corresponde à variável classificação na prova final. Com esses pares ordenados, representamos o diagrama de dispersão seguinte.



Fonte: Própria

47.2. A variável independente é a classificação interna.

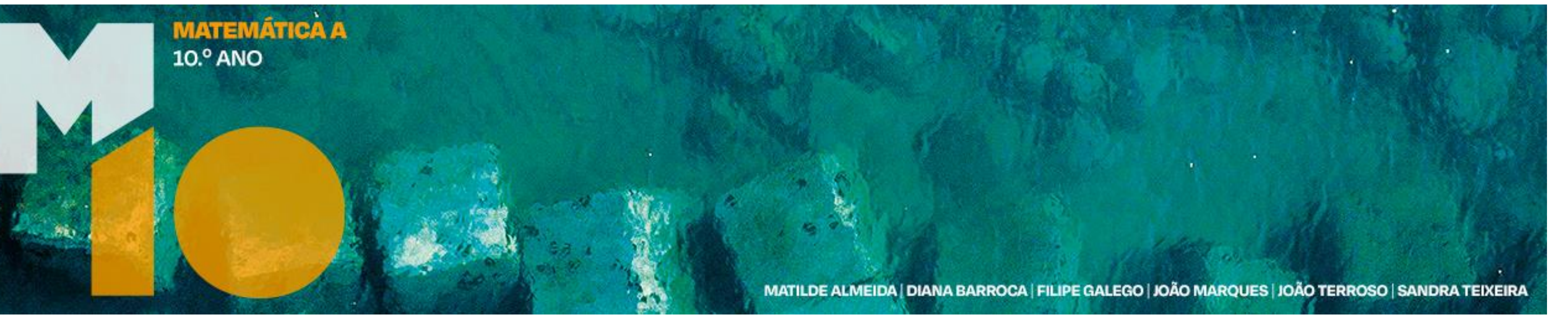
49.3. A variável dependente é a classificação na prova final de Português.

49.4. Do conjunto de 9 alunos do 9.º A, aqueles que tiveram melhores resultados na classificação interna, também tiveram, tendencialmente, melhores resultados na prova final de Português.

Aplicar - **Página 148**

50. Por observação de cada um dos diagramas de dispersão representados, podemos fazer as correspondências seguintes.

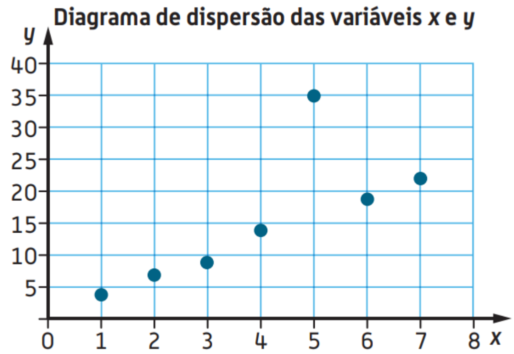
(A) – (II) (B) – (III) (C) – (V) (D) – (I)



Aplicar - Página 149

49.

49.1. Considerando os pares ordenados (x_i, y_i), conseguimos construir o diagrama de dispersão seguinte.



Fonte: Própria

49.2. O ponto $(5, 35)$ assinalado a vermelho não segue o padrão dos restantes. Pode ser considerado um *outlier*.

49.3. No caso de incluirmos o *outlier* no conjunto de pares ordenados, o coeficiente de correlação é, aproximadamente, 0,752.

49.4. Ao excluirmos o *outlier* do conjunto de pares ordenados, o coeficiente de correlação é 0,996.

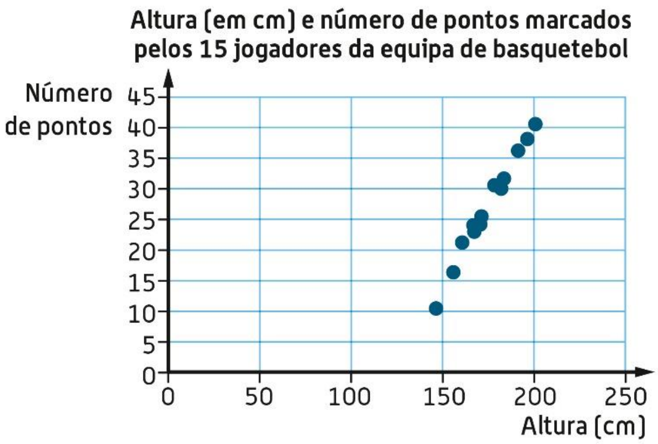
49.5. Se não se incluir o *outlier* (5, 35), a correlação entre as duas variáveis passa a ser muito forte, como seria de esperar.

Aplicar - Página 150

50.

Seja o conjunto de pares ordenados (x_i, y_i), em que x_i é a variável altura (em centímetros dos jogadores do plantel e y_i é a variável número de pontos marcados nos 5 treinos da semana. Com estes pares ordenados conseguimos representar o diagrama de dispersão seguinte.

Observa-se uma forte correlação positiva entre as variáveis, corroborada pelo valor aproximado do coeficiente de correlação $r \approx 0{,}995$.



Fonte: Própria

51. Comecemos por escrever os dados em duas colunas de uma folha de cálculo.

	A	B
1	Idade	Nível de Glicose
2	42	96
3	20	54
4	24	60
5	44	75
6	57	87
7	59	89

Se pretendermos determinar o coeficiente de correlação utilizamos a fórmula =CORREL(A2:A7;B2:B7) na folha de cálculo.

	A	B	C	D	E
1	Idade	Nível de Glicose			
2	42	96			
3	20	54			
4	24	60		Coeficiente	=CORREL(A2:A7;B2:B7)
5	44	75			
6	57	87			
7	59	89			

	A	B	C	D	E
1	Idade	Nível de Glicose			
2	42	96			
3	20	54			
4	24	60		Coeficiente	0,84597715
5	44	75			
6	57	87			
7	59	89			

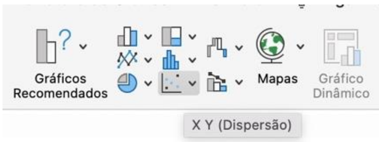
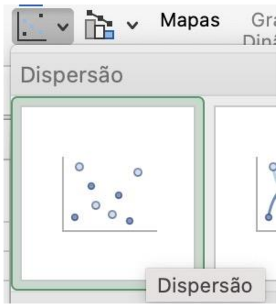
O coeficiente de correlação é, aproximadamente, 0,846.

Tarefa 16 - Página 151

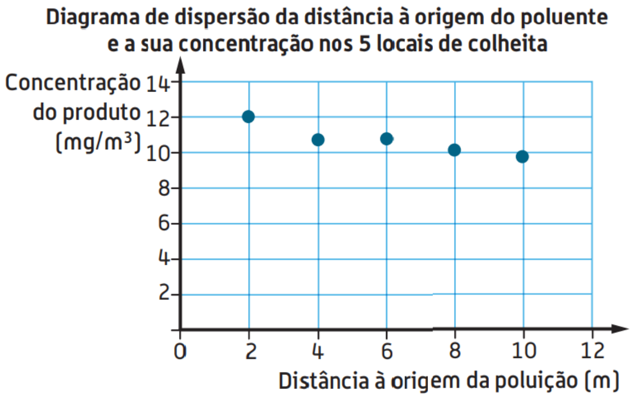
1. Começamos por definir duas listas numa folha de cálculo tal como se mostra a seguir.

	A	B
1	Distância à origem da poluição	Concentração do produto poluente
2	2	12
3	4	10,7
4	6	10,8
5	8	10,2
6	10	9,8

Considerando os pares ordenados (x_i, y_i), em que x_i corresponde à variável distância à origem da poluição e y_i corresponde à variável concentração do produto poluente, seguimos os seguintes passos para representar o diagrama de dispersão: selecionamos as células A2 a A6 e B2 a B6; selecionamos Inserir, seguido de X Y (dispersão); por último, selecionamos Dispersão.

1.º passo	2.º passo	3.º passo
		

Desta forma, obtemos o diagrama de dispersão seguinte.



Fonte: Própria

2.

2.1. Para obter o coeficiente de correlação, numa célula, escrevemos a fórmula $=CORREL(A2:A6;B2:B6)$.

O coeficiente de correlação é, aproximadamente, $-0{,}933$.

Podemos observar que existe uma correlação linear forte.

2.2. A reta de regressão pode ser obtida seguindo os passos seguintes.

1.º passo

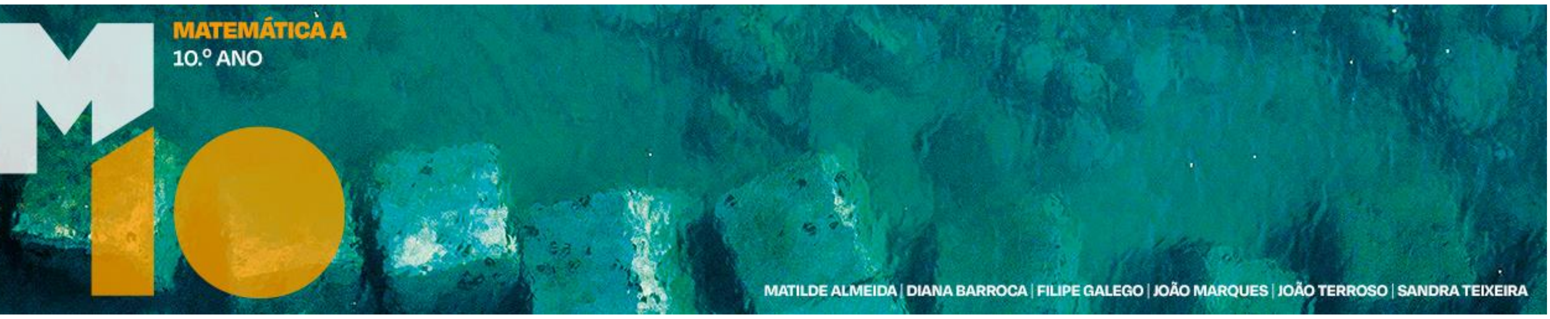
Selecionamos a representação gráfica do diagrama de dispersão.

2.º passo

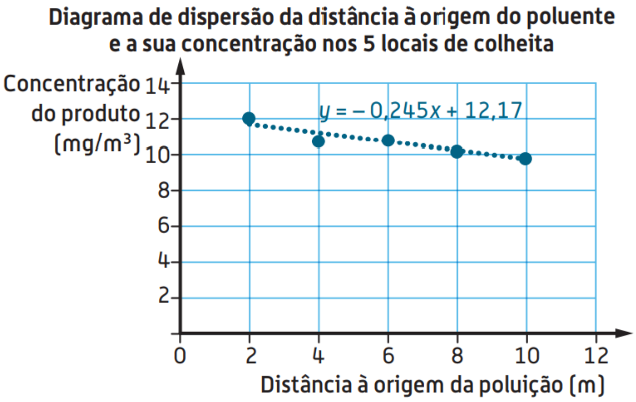
Selecionamos Estrutura do Gráfico, seguida de Esquema rápido.

3.º passo

Selecionamos Esquema 9 (ou equivalente).



A reta regressão fica representada da forma seguinte.



Fonte: Própria

A equação da reta de regressão é $y = -0{,}245x + 12{,}17$.

3. Consideremos a equação da reta de regressão obtida anteriormente.

$$y = -0{,}245x + 12{,}17$$

Usando esta equação, determinamos um valor aproximado da concentração do produto poluente, sabendo que $x = 20$ m.

$$y = -0{,}245 \times 20 + 12{,}17 = 7{,}27$$

O valor da concentração do produto poluente quando a distância à origem da poluição for igual a $20\ m$ é, aproximadamente, $7{,}27\ \mathrm{mg/m^3}$.

4. Para que a poluição não ser detetável, então a concentração do produto poluente é igual a 0.

Considerando a equação da reta de regressão e tendo em atenção que $y = 0$, temos que:

$$0 = -0{,}245x + 12{,}17 \Leftrightarrow 0{,}245x = 12{,}17 \Leftrightarrow x = \frac{12{,}17}{0{,}245} \Leftrightarrow x \approx 49{,}67$$

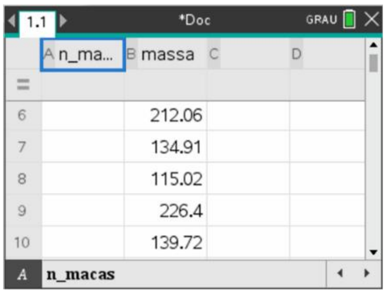
Para que a poluição já não seja detetável é necessário estar a, aproximadamente, $49{,}67\ m$ da origem da mesma.

Aplicar - Página 152

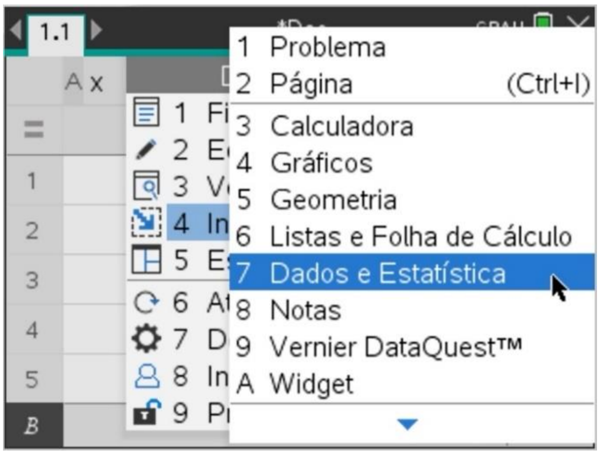
52.

52.1.

Recorrendo às potencialidades da calculadora (Ti-Nspire), começamos por colocar os dados em duas listas numa folha de cálculo (dados e folha de cálculo).

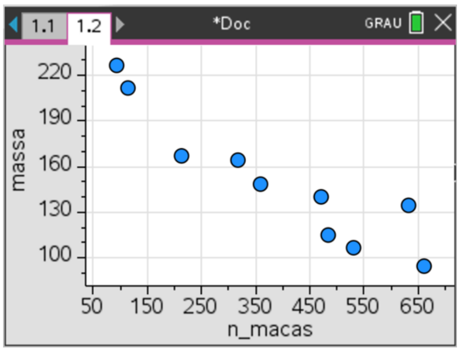


De seguida, carregando em doc, escolhemos a opção *4 Inserir*, seguida da opção *7 Dados e Estatística*.

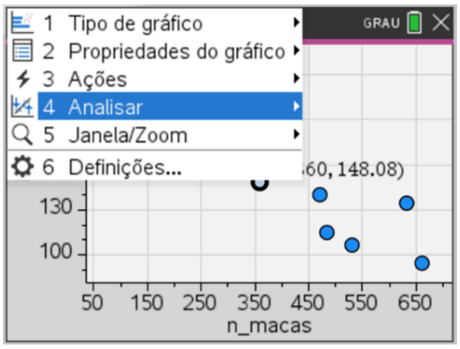


Na janela seguinte, iremos definir os eixos. No eixo das abcissas, selecionamos o número de maçãs e, no eixo das ordenadas, selecionamos a massa.

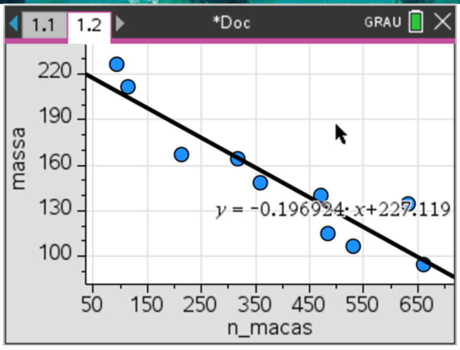
Desta forma, obtemos o diagrama de dispersão seguinte.



52.2. Carregamos na tecla MENU e em *4 Analisar*.



Selecionamos a opção *6 Regressão*, seguida da opção *1 Mostrar linear* $(mx + b)$.

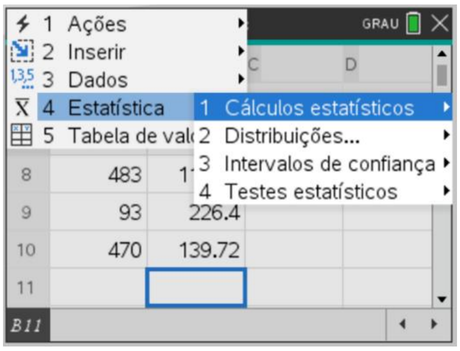


Obtemos a equação da reta de regressão seguinte.

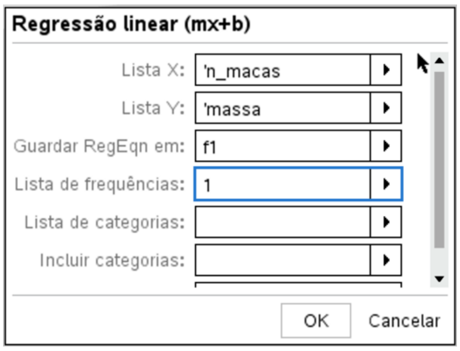
$$y = -0{,}1969237x + 227{,}119$$

Caso pretendamos apenas a equação, podemos não traçar o diagrama de dispersão e obter apenas a equação da reta.

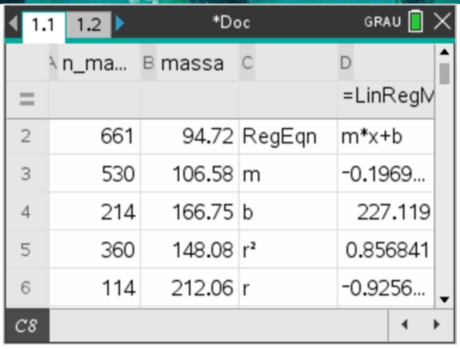
Para isso, carregamos em MENU, de seguida *4 Estatística*. A seguir, selecionamos a opção *1 Cálculos estatísticos* e *3 Regressão linear* $(mx + b)$.



Em seguida, identificamos as listas onde inserimos os dados (*Lista X* corresponde ao número de maçãs e *Lista Y* corresponde à massa).



Desta forma, obtemos os dados da equação da reta na tabela.



	A n_ma...	B massa	C	D
=				=LinRegM
2	661	94.72	RegEqn	m*x+b
3	530	106.58	m	-0.1969...
4	214	166.75	b	227.119
5	360	148.08	r²	0.856841
6	114	212.06	r	-0.9256...

$y = -0{,}1969237x + 227{,}119$

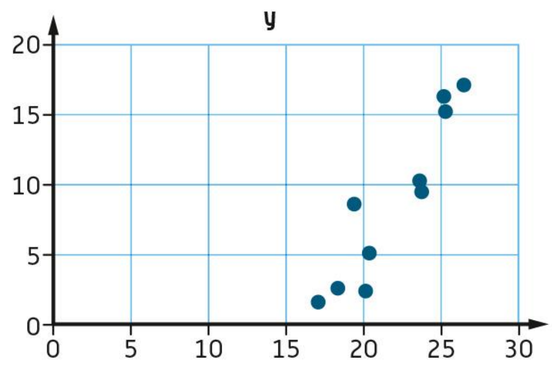
52.3. Atendendo ao valor do coeficiente de correlação $r \approx -0{,}926$, podemos afirmar que se trata de uma correlação negativa forte.

Aplicar - Página 154

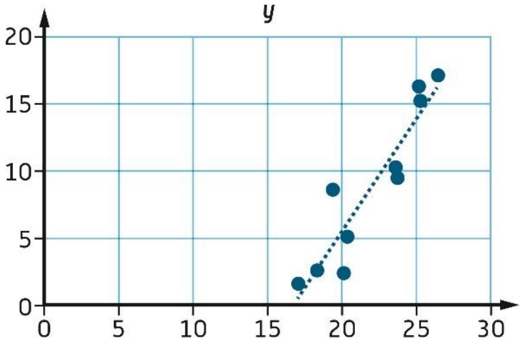
53. Comecemos por introduzir os dados numa folha de cálculo.

	A	B
1	x	y
2	17	1,5
3	18,3	2,5
4	19,4	8,5
5	20,1	2,3
6	20,3	5,1
7	23,6	10
8	23,7	9,5
9	25,1	16,1
10	25,2	15,1
11	26,2	17

Para determinarmos a equação da reta de regressão linear, vamos começar por construir o diagrama de dispersão.



De seguida, acrescentamos a reta de regressão.



A equação da reta de regressão, com os parâmetros arredondados às milésimas, é

$$y = 1{,}683x - 28{,}088$$

Obtemos os valores seguintes.

$$a \approx 1{,}683 \qquad b \approx -28{,}088$$

Para estimarmos, com base neste modelo, o tempo que um turista desse grupo com massa corporal 23,1 demora a subir aquela escadaria, fazemos $x = 23{,}1$.

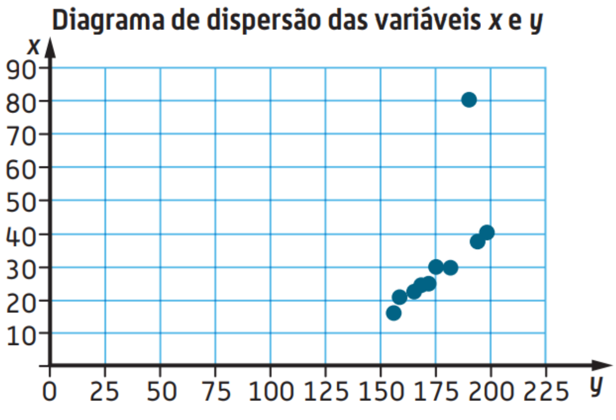
$$y = 1{,}683 \times 23{,}1 - 28{,}088 \approx 10{,}8$$

Assim, o tempo de subida para um turista com IMC de 23,1 é estimado em, aproximadamente, 10,8 minutos.

Aplicar - Página 155

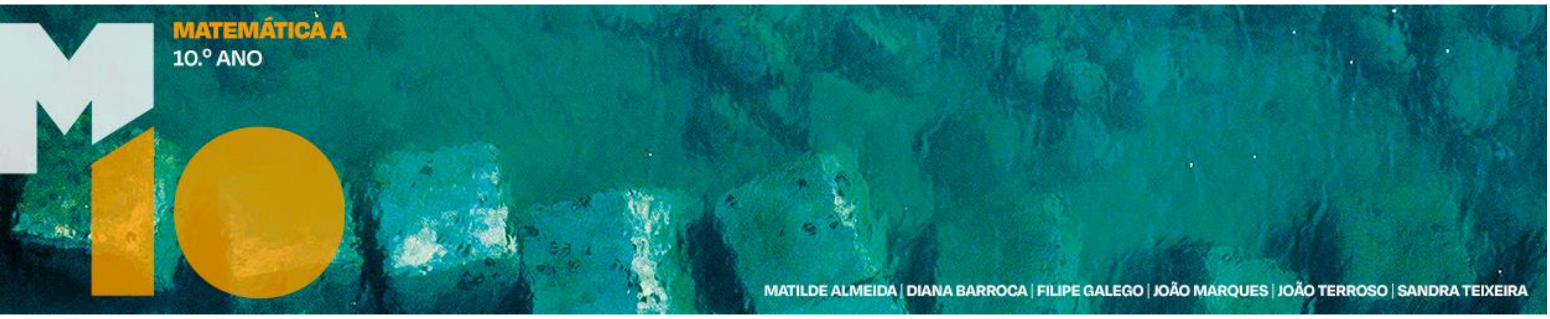
54.

54.1. Seja (x_i, y_i) o par ordenado em que a x_i correspondem os valores assumidos pela variável x e a y_i correspondem os valores assumidos pela variável y.

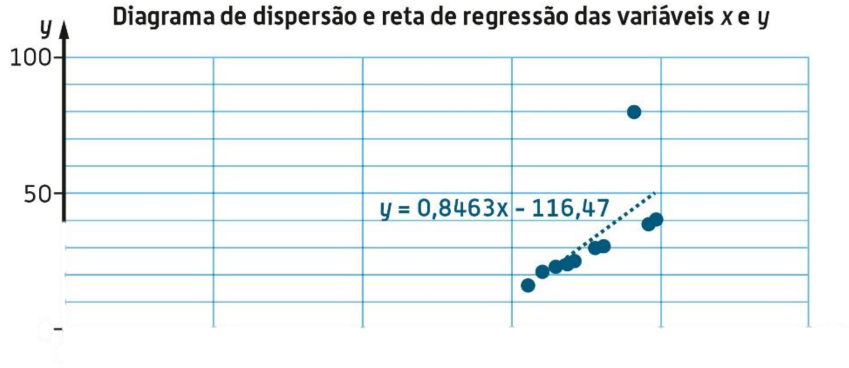


Fonte: Própria

Com este conjunto de pares ordenados, obtemos o diagrama de dispersão seguinte.



54.2. Usando uma folha de cálculo, obtemos a representação gráfica seguintes.

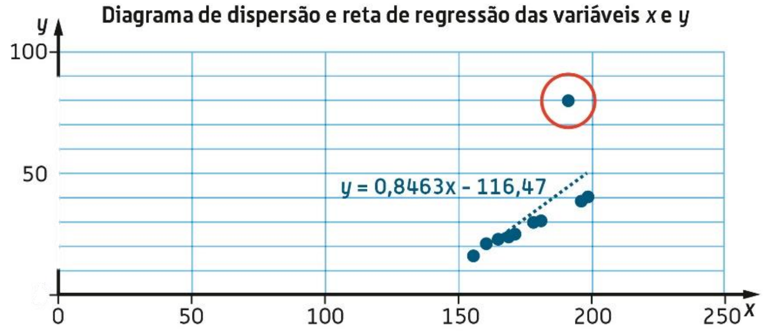


Fonte: Própria

A equação da reta de regressão é $y = 0{,}8463x - 116{,}47$.

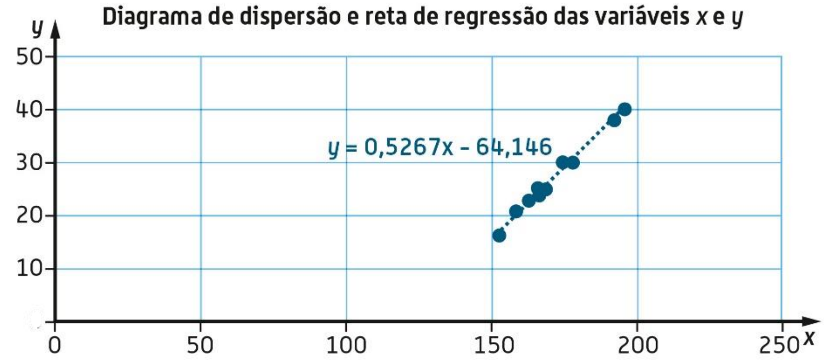
54.3. Recorrendo às potencialidades de uma folha de cálculo, o coeficiente de correção é, aproximadamente, $r \approx 0{,}693$.

54.4. Observando a representação gráfica do diagrama de dispersão e da reta de regressão, conseguimos identificar o *outlier* $(190{,}80)$ (assinalado em baixo).



Fonte: Própria

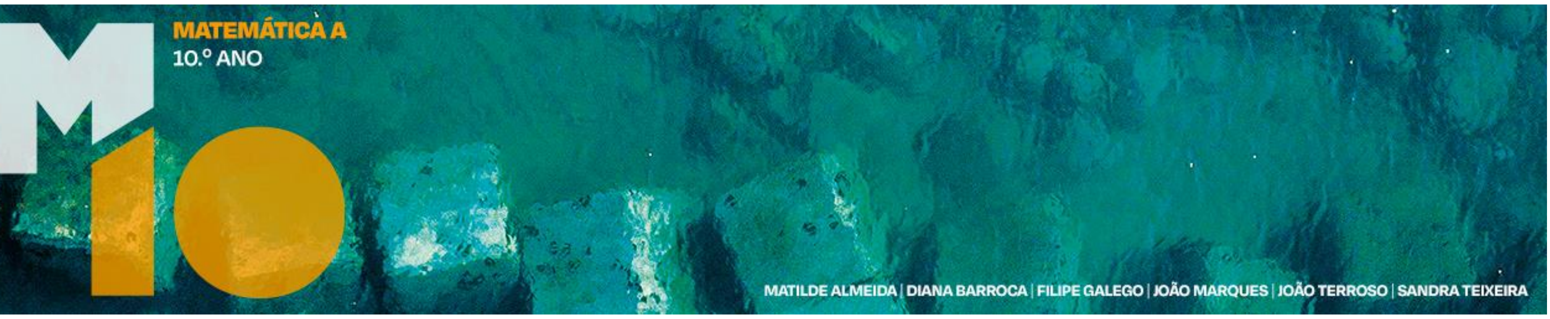
54.5. Retirando o *outlier* $(190{,}80)$ do conjunto de pares ordenados, obtemos as seguintes representações do diagrama de dispersão e da reta de regressão.



Fonte: Própria

A reta de regressão é definida pela equação $y = 0{,}5267x - 64{,}146$, sendo o coeficiente de correlação $r \approx 0{,}994$.

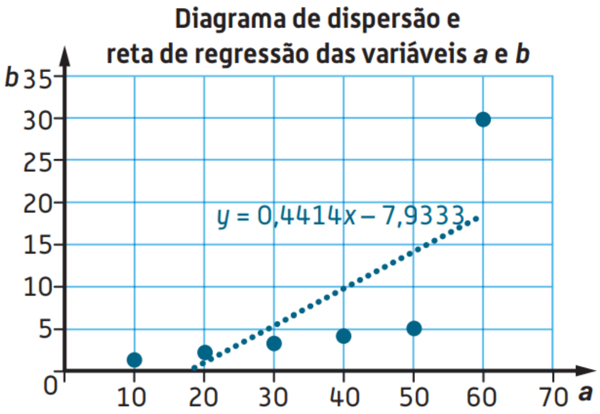
54.6. Ao retirar-se o *outlier* $(190{,}80)$ do conjunto de pares ordenados, a correlação tornou-se mais forte e os pontos do diagrama de dispersão ficaram mais próximos da reta de regressão.



Aplicar - Página 156

55.

55.1. Considerando o conjunto de pares ordenados (x_i, y_i), em que x_i corresponde aos valores assumidos pela variável a e y_i corresponde aos valores assumidos pela variável b, obtemos o diagrama de dispersão e a reta de regressão seguintes.



Fonte: Própria

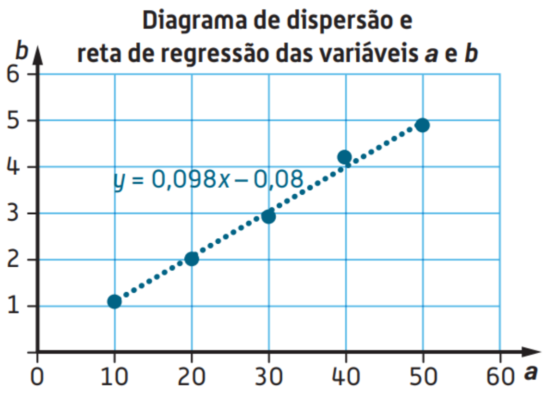
Temos que:

$$r \approx 0{,}744 \qquad a \approx 0{,}441 \qquad b \approx -7{,}933$$

A reta de regressão é definida pela equação $y = 0{,}441x - 7{,}933$ e o valor do coeficiente de correlação é, aproximadamente, $r \approx 0{,}744$.

55.2. Observando o diagrama de dispersão, podemos afirmar que o *outlier* é o par ordenado $(60{,}30)$.

55.3. Retirando o *outlier* $(60{,}10)$ do conjunto de pares ordenados, obtemos os diagrama de dispersão e reta de regressão seguintes.

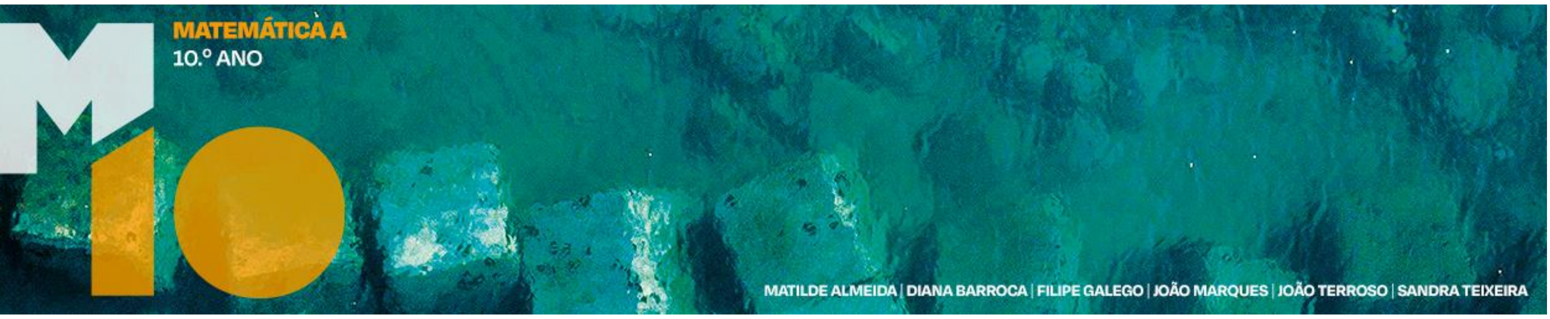


Fonte: Própria

Neste caso:

$$r \approx 0{,}997 \qquad a = 0{,}098 \qquad b = 0{,}08$$

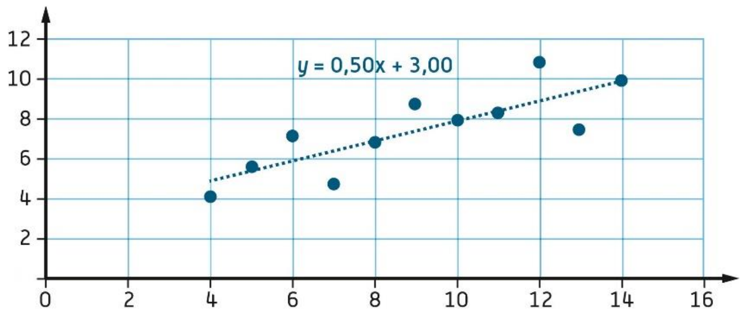
A nova reta de regressão é definida pela equação $y = 0{,}098x + 0{,}08$, sendo o coeficiente de correlação $r \approx 0{,}997$.



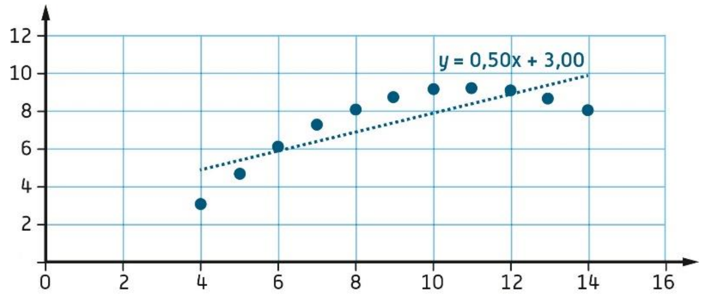
Aplicar - Página 157

56. Com recurso à folha de cálculo, obtemos os diagramas de dispersão seguintes.

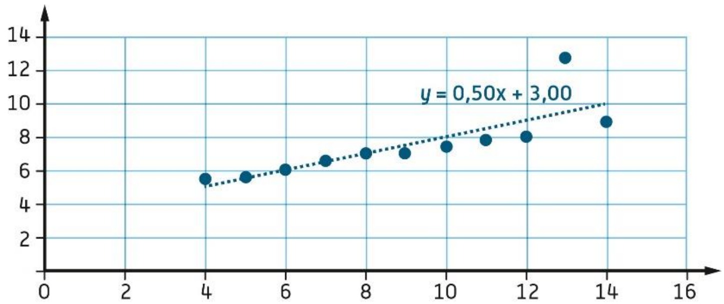
I



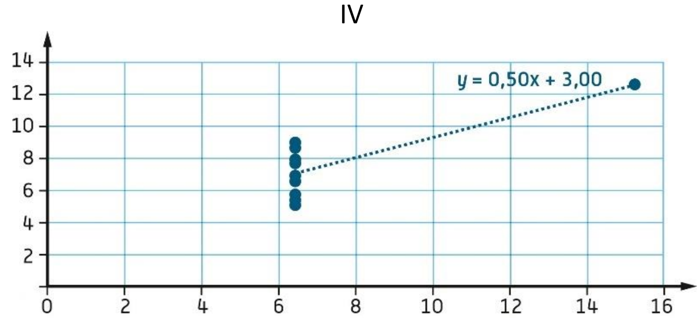
II



III

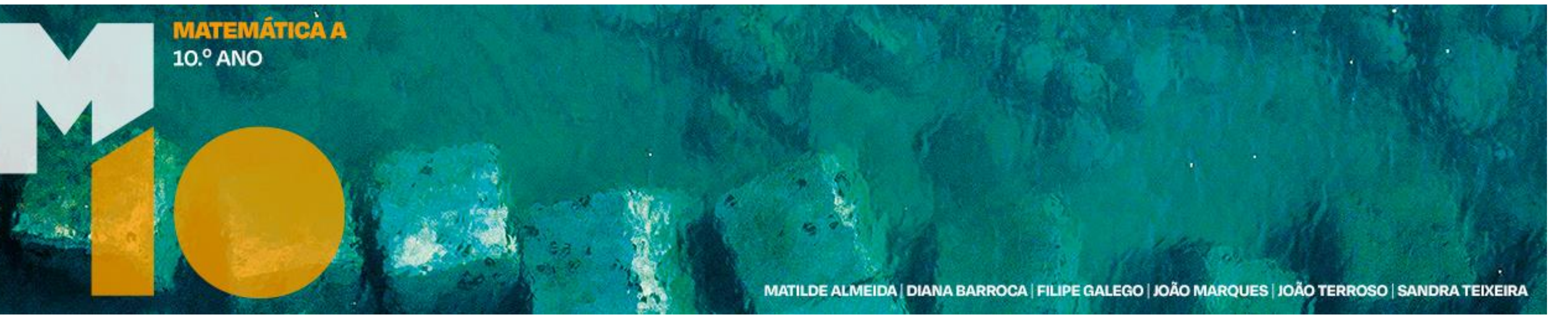


IV



Nos quatro casos, $r = 0{,}816$ e a equação da reta de regressão é $y = 0{,}50x + 3{,}00$

A correlação linear não é a mais adequada nos casos II e IV.
Observam-se *outliers* em III e IV.



57. Por observação de cada um dos diagramas de dispersão representados, podemos fazer as correspondências seguintes.

(A) – (ii) (B) – (i) (C) – (iii) (D) – (v) (E) – (iv)

Aplicar - Página 158

58.

58.1. Uma explicação possível seria a de que, num momento de crescimento económico, o consumo tem tendência a aumentar, seja de carne, de carros ou de outros produtos.

58.2. Uma população que possui muitos eletrodomésticos é uma população com um nível de vida razoavelmente alto, logo, com melhores condições de vida e melhores cuidados de saúde, portanto, com maior esperança de vida.

58.3. O aumento dos pés de uma criança está relacionado com a sua idade e quanto mais velha ela é, mais terá aprendido, aumentando os seus conhecimentos e a sua capacidade de resolver problemas.

58.4. Quanto mais sol houver, maior é necessidade de comprar óculos de sol e maior é a vontade de comer gelados, pois o calor também aumenta.

Tarefa 17 – Página 159

1. A temperatura mínima atingida neste mês foi 5°C. Esta temperatura foi atingida no dia 23.

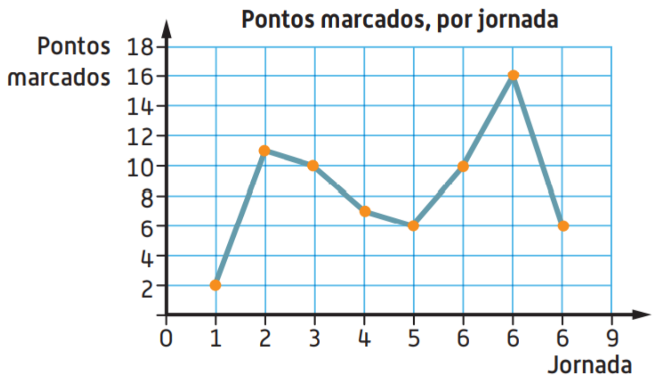
2. A temperatura máxima foi 20°C em dois dias. Nos dias 12 e 19.

3. Existe um aumento da temperatura máxima entre os dias 17 e 19 de novembro (por exemplo) e uma diminuição da temperatura máxima entre os dias 4 e 6 de novembro (por exemplo).

Aplicar – Página 159

59.

Consideremos o conjunto de pares ordenados (x_i, y_i), em que a x_i corresponde à variável jornada e y_i corresponde à variável pontos marcados pelo jogador, conseguimos construir o seguinte gráfico de linhas.



Fonte: Própria

Ao longo destas 8 jornadas, a prestação do jogador mostrou ser irregular.
Na primeira jornada, o jogador regista o menor número de pontos marcados no conjunto das 8 jornadas.

Da primeira para a segunda jornada, verificou-se um aumentou do número de pontos marcados. No entanto, durante as 3 jornadas seguintes, esse número de pontos foi diminuindo.

Nas sexta e sétima jornadas, o jogador voltou a melhorar o número de pontos marcados, tendo atingido o seu valor máximo na sétima jornada.

Na oitava jornada, o número de pontos marcados voltou a diminuir.

Aplicar +
Páginas 164 a 167

Itens de seleção
1. Opção (D)

2. Tendo em conta a variedade de marcas à venda de batatas fritas no mercado português, não se justifica estudar a população. Pelo que deveremos selecionar uma amostra.

Opção (B)

3. A amostra é composta por 450 espetadores do jogo de basquetebol.

Opção (B)

4. Uma amostra representativa para esta pesquisa seria selecionar 200 jovens que passeavam na rua.

Opção (D)

5. Temos que:

- $60\% \times 360 = 0{,}6 \times 360 = 216$

Opção (C)

6.
6.1. A variável corresponde ao tempo, em horas, de visualização dos programas do canal de televisão. Logo, é uma variável quantitativa contínua.

Opção (D)

6.2. O dado da Rita está inserido na classe $[6{,}8[$.

Opção (D)

6.3. O número de espetadores que respondeu ao inquérito foi:

$$N.º\ de\ espetadores = 40 + 30 + 15 + 10 = 95$$

A média de horas é:

$$\bar{x} = \frac{1 \times 40 + 3 \times 30 + 5 \times 15 + 7 \times 10}{95} \approx 3$$

Opção (C)

7. Temos que:

- I é verdadeira, pois, no diagrama 3, tem-se que $mínimo = 1 = \tilde{x} - 2{,}5$; $máximo = 4{,}5 = \tilde{x} + 1{,}5$
- II é falsa. Na distribuição 2, o mínimo é 1.
- III é falsa. Na distribuição 1, pelo menos 50% de dados estatísticos são maiores ou iguais a 4,5.
- IV é falsa. A amplitude interquartil da distribuição de dados associada ao diagrama 3 é 2, pois, $Q_3 - Q_1 = 4{,}5 - 2{,}5 = 2$.
- V é verdadeira. 50% dos dados pertencem ao intervalo [1, 2], que tem amplitude 1, e outros 50% dos dados pertencem ao intervalo [2, 6], que tem amplitude 4, logo a afirmação é verdadeira.

Opção (A)

8.

8.1. A amplitude das massas do 10.º G é:

$$Amplitude = máx. - mín. = 80 - 50 = 30$$

A amplitude das massas dos alunos do 10.º G é 30 kg.

Opção (C)

8.2. A moda é o valor que assuma a maior frequência absoluta no conjunto das massas do 10.º G. Podemos então concluir que $M_o = 56$.

Opção (B)

8.3. Foram coletados 24 dados. A mediana irá corresponder à média aritmética das massas que ocupam os 12.º e 13.º lugares, aquando da ordenação das massas da turma por ordem crescente.

$$M_e = \frac{x_{12} + x_{13}}{2} = \frac{49 + 50}{2} = 49{,}5$$

O conjunto de dados fica então dividido em dois grupos: o primeiro grupo composto pelos dados dos alunos das posições 1 a 12; o segundo grupo composto pelos dados dos alunos das posições 13 a 24.

Os valores do 1.º e 3.º quartil são:

$$Q_1 = \frac{a_6 + a_7}{2} = \frac{55 + 56}{2} = 55{,}5 \qquad Q_3 = \frac{a_{17} + a_{18}}{2} = \frac{63 + 64}{2} = 63{,}5$$

Concluímos que $Q_1 = 55{,}5$ e $Q_3 = 63{,}5$.

Opção (A)

9.

9.1. A média da variável x é:

$$\bar{x} = \frac{1 \times 9 + 2 \times 2 + 3 \times 6 + 4 \times 6 + 5 \times 7}{30} = \frac{90}{30} = 3$$

Opção (A)

9.2. Existem 30 dados estatísticos. Ao colocar os dados ordenados, a mediana corresponderá à média entre os 15.º e 16.º dados, ou seja:

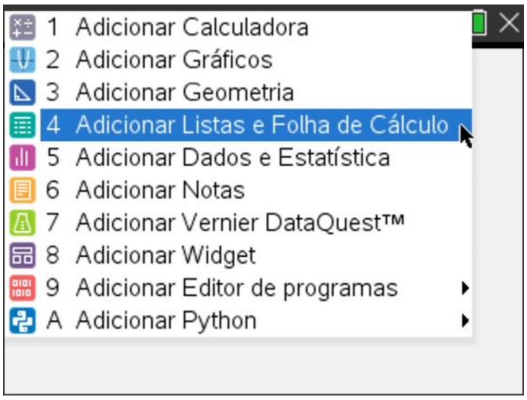
$$M_e = \frac{3+3}{2} = 3$$

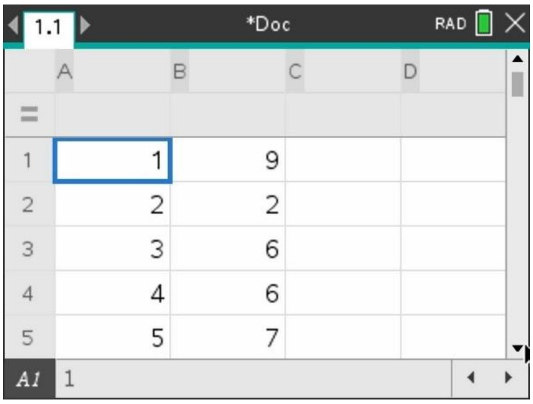
Opção (B)

9.3. Recorrendo às potencialidades da calculadora gráfica, percorremos as seguintes etapas:

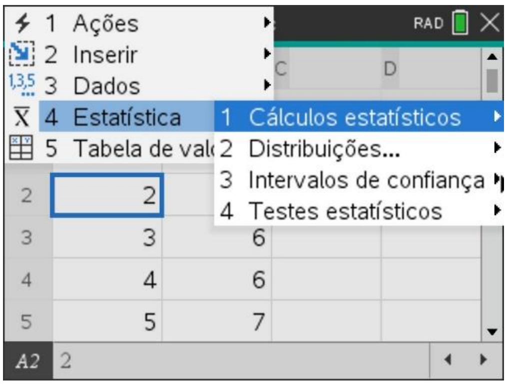
- *TI-Nspire*

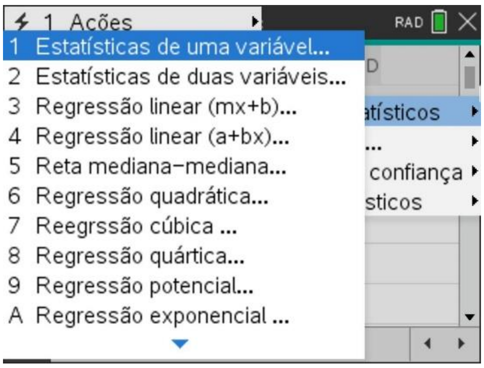
Começamos por selecionar *Listas e Folha de Cálculo*. De seguida, nas colunas *A* e *B* introduzimos os valores da variável x e as respetivas frequências absolutas.



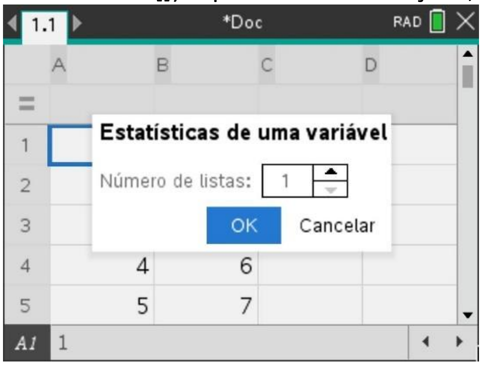


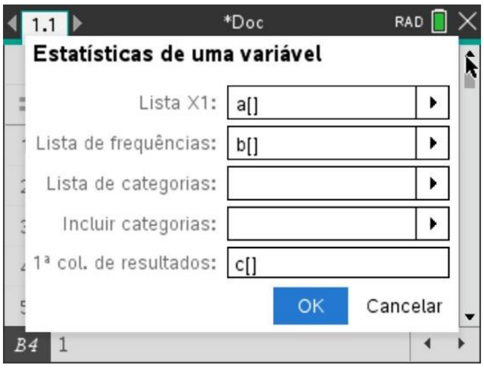
Após introduzidos os dados, selecionamos menu, seguido de *4 Estatística*, *1 Cálculos estatísticos* e *1 Estatística de uma variável*.



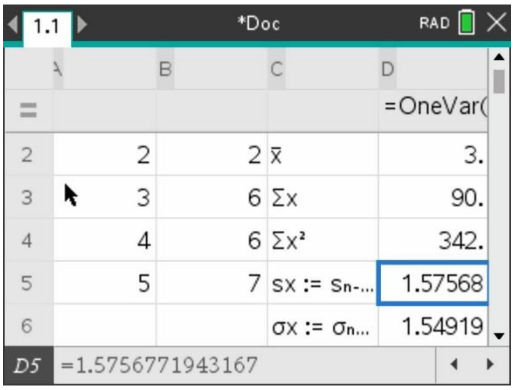


Preenchemos o números de listas a considerar (neste caso, 1), premimos OK, e verificamos as listas que a calculadora deverá considerar (Lista X1 : a[], Lista de frequências: b[] e 1.º col. De resultados: c[]). Após estas verificações, premimos OK.



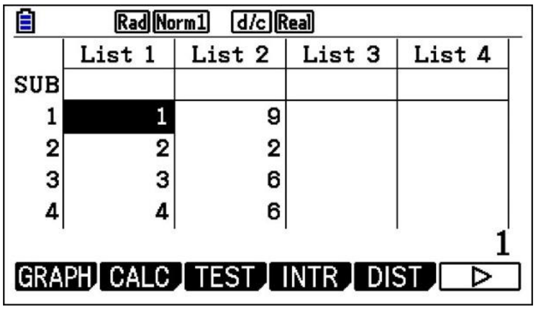


A calculadora devolve-nos os valores seguintes.

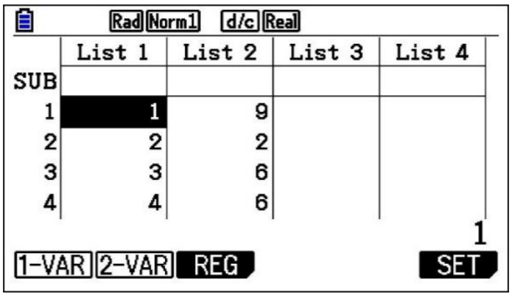


- *CASIO fx-CG50*

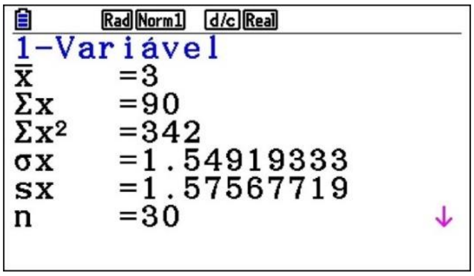
Começamos por selecionar *Estatística*. De seguida, nas colunas *Lista 1* e *Lista 2* introduzimos os valores da variável x e as respetivas frequências absolutas.



Após a introdução dos dados, selecionamos *F2 CALC* e procedemos à verificação das definições em *F6 SET*. Definimos *1Var XList : Lista 1*. Em *1Var Freq:*, carregamos em *F2 LIST*, sendo que introduzimos *2*. Na nossa janela, surge então *1Var XList : Lista 1* e *1Var Freq: Lista 2*.

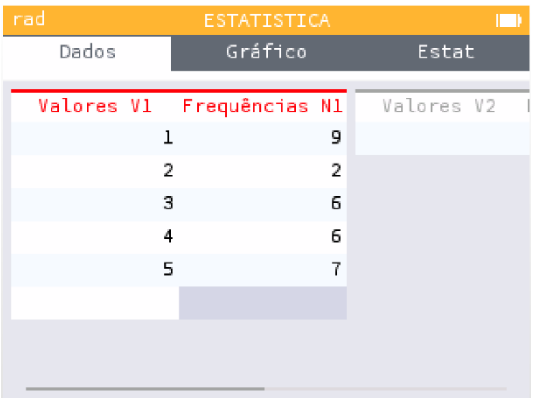


Fazendo *EXIT*, e selecionando *F1 1-VAR*, a máquina devolve a janela de valores seguinte.

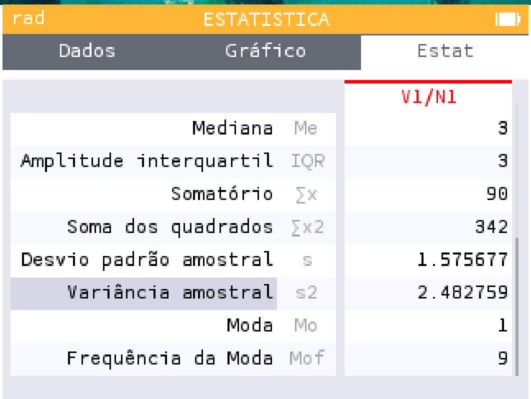


- *NumWorks*

Começamos a selecionar *Estatística*. De seguida, introduzimos os dados nas listas *Valores V1* e *Frequências N1* dos valores da variável x e respetivas frequências absolutas.



De seguida, selecionamos *Estat* e a máquina devolve uma janela com o conjunto de valores seguintes.



		V1/N1
Mediana	Me	3
Amplitude interquartil	IQR	3
Somatório	∑x	90
Soma dos quadrados	∑x2	342
Desvio padrão amostral	s	1.575677
Variância amostral	s2	2.482759
Moda	Mo	1
Frequência da Moda	Mof	9

Concluímos que o valor do desvio padrão é, aproximadamente, 1,58.

Opção (C)

9.4. Temos que:

- $Frequência\ relativa\ 3 = \frac{6}{30} = 0{,}2$

Concluímos que a frequência relativa de 3 é 0,2.

Opção (A)

10. Com a situação de saúde do Gilberto, temos que a média diminuiu 2 valores, mas o desvio padrão mantém-se, pois, todas as classificações baixaram o mesmo número de valores.

$$\bar{x} = 17{,}2 - 2 = 15{,}2 \qquad s^2 = 2{,}2$$

Opção (C)

11. Por hipótese, a o valor da média das idades dos alunos é 15,4. Assim:

$$\bar{x} = 15{,}4 \Leftrightarrow \frac{14 \times 6 + 15 \times 5 + 16 \times a + 17 \times 5}{16 + a} = 15{,}4 \Leftrightarrow$$

$$\Leftrightarrow 244 + 16a = (16 + a) \times 15{,}4 \Leftrightarrow$$

$$\Leftrightarrow 244 + 16a = 246{,}4 + 15{,}4a \Leftrightarrow$$

$$\Leftrightarrow 16a - 15{,}4a = 246{,}4 - 244 \Leftrightarrow$$

$$\Leftrightarrow 0{,}6a = 2{,}4 \Leftrightarrow$$

$$\Leftrightarrow a = 4$$

Opção (B)

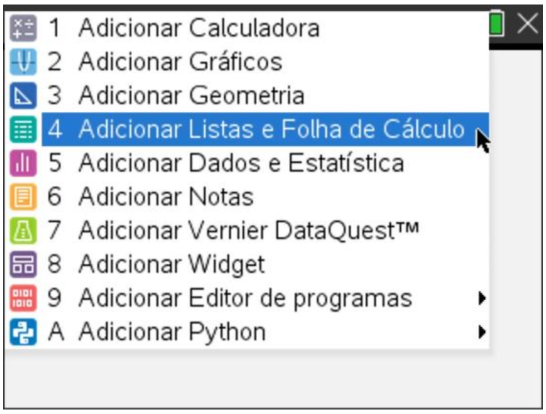
12. Começamos por organizar os dados numa tabela de frequências absolutas simples.

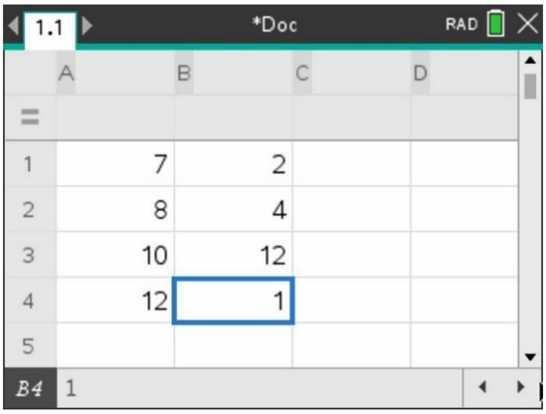
x	Frequência absoluta
7	2
8	4
10	12
12	1

Recorrendo às potencialidades da calculadora gráfica, percorremos as seguintes etapas:

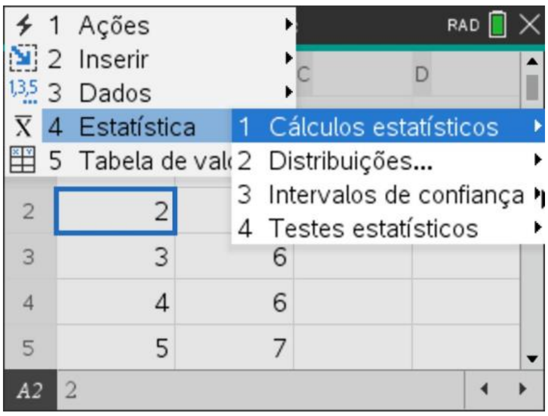
- *TI-Nspire*

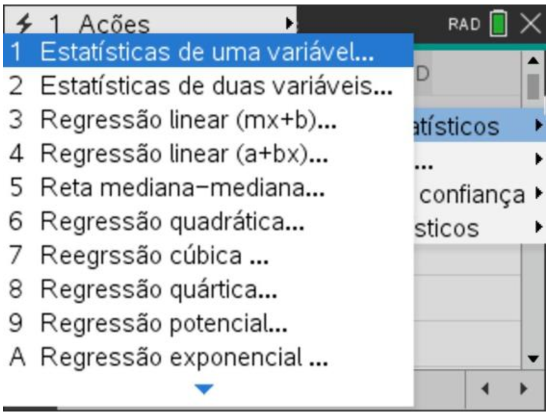
Começamos por selecionar *Listas e Folha de Cálculo*. De seguida, nas colunas *A* e *B* introduzimos os valores da variável x e as respetivas frequências absolutas.



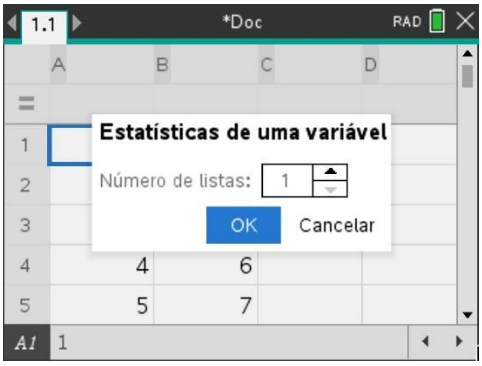


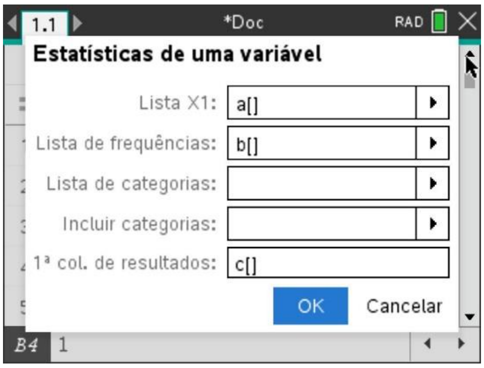
Após introduzidos os dados, selecionamos menu, seguido de *4 Estatística*, *1 Cálculos estatísticos* e *1 Estatística de uma variável*.



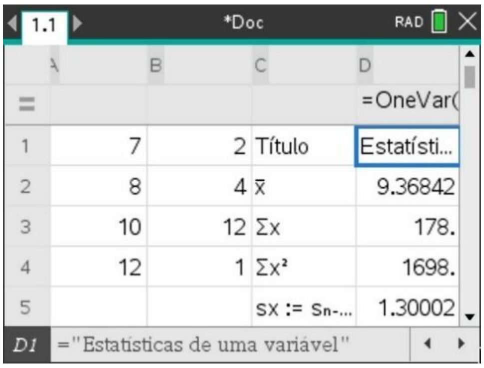


Preenchemos o números de listas a considerar (neste caso, 1), premimos OK, e verificamos as listas que a calculadora deverá considerar (Lista X1 : a[], Lista de frequências: b[] e 1.º col. De resultados: c[]). Após estas verificações, premimos OK.



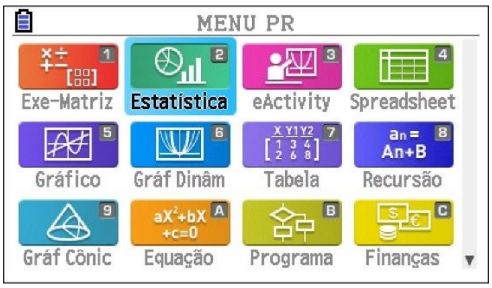


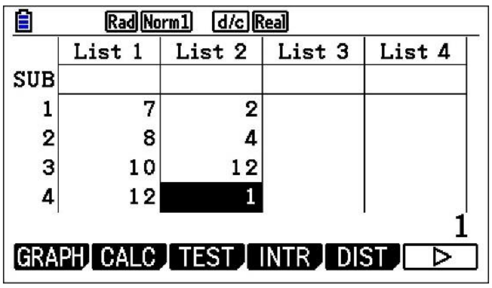
A calculadora devolve-nos os valores seguintes.



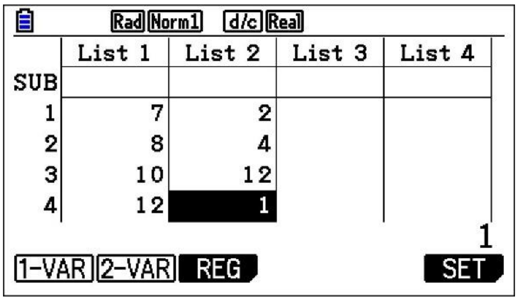
- *CASIO fx-CG50*

Começamos por selecionar *Estatística*. De seguida, nas colunas *Lista 1* e *Lista 2* introduzimos os valores da variável x e as respetivas frequências absolutas.

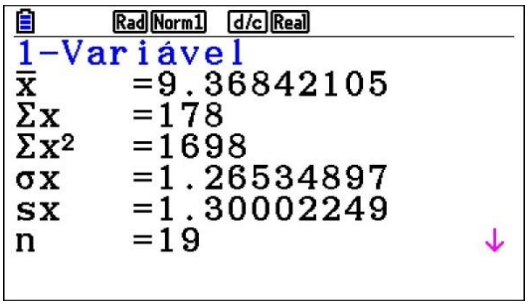




Após a introdução dos dados, selecionamos *F2 CALC* e procedemos à verificação das definições em *F6 SET*. Definimos *1Var XList : Lista 1*. Em *1Var Freq:*, carregamos em *F2 LIST*, sendo que introduzimos *2*. Na nossa janela, surge então *1Var XList : Lista 1* e *1Var Freq: Lista 2*.



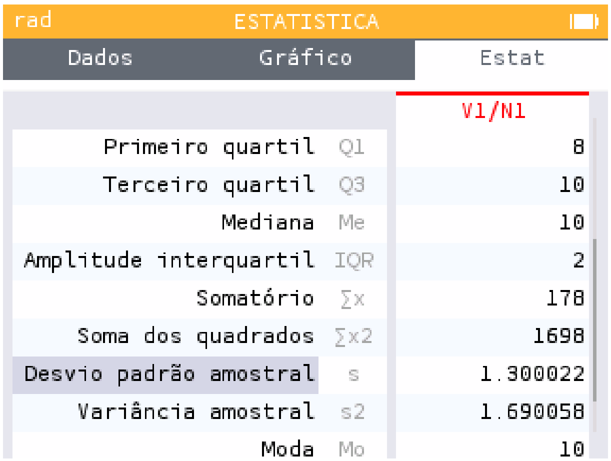
Fazendo *EXIT*, e selecionando *F1 1-VAR*, a máquina devolve a janela de valores seguinte.



- *NumWorks*

Começamos a selecionar *Estatística*. De seguida, introduzimos os dados nas listas *Valores V1* e *Frequências N1* dos valores da variável x e respetivas frequências absolutas.

De seguida, selecionamos *Estat* e a máquina devolve uma janela com o conjunto de valores seguintes.



Concluímos que o valor do desvio padrão é, aproximadamente, 1,3.

Opção (C)

13. Como se trata de uma correlação positiva forte, então, $r > 0$ e próxima de 1. Excluímos as opções (A) e (B).
Como o declive da reta de regressão é positivo, então, excluímos a opção (C).
A única opção que verifica ambas as condições é a (D).

Opção (D)

14. À medida que a variável x aumenta, a variável y também aumenta. Logo, trata-se de uma correlação positiva. Temos então que $r > 0$. Excluímos as opções (B) e (D).
Como o declive, a ,da reta de regressão é positivo, então, excluímos a opção (C).
A única resposta que verifica ambas as condições é a (A).

Opção (A)

Aplicar +
Páginas 168 - 177

Itens de construção

15.
15.1. Temos que:

A. A população em estudo é o conjunto dos ouvintes da rádio referida.

B. A população em estudo é população portuguesa.

C. A população em estudo é o conjunto dos alunos das 5 turmas da escola referida.

15.2. No caso A, a dimensão da amostra é 100. No caso B, a amostra tem dimensão igual a 220. No caso C, a amostra é constituída por 175 elementos.

15.3. Apenas na situação C a amostra escolhida é representativa da população. Na situação A. foi usada uma amostra por resposta voluntária. Logo, podem ter telefonado apenas os que gostam de certo programa. Na situação B. foi usada uma amostra por conveniência, mas, quem vai a um Centro de Saúde, em princípio, será quem não se automedica.

16.
16.1. Temos que, a população em estudo é:
(I) Conjunto dos trabalhadores da empresa Fonte Nova.
(II) Conjunto das baterias para telemóvel de uma certa marca.
(III) Conjunto dos alunos do 9.º ano de uma turma.
(IV) População portuguesa.
(V) Conjunto dos refrigerantes vendidos por uma certa empresa.

16.2. Tendo em conta o estudo que se pretende fazer e a respetiva população, podemos afirmar que em (I) e (III) pretende-se realizar um censo e em (II), (IV) e (V) pretende-se realizar uma sondagem.

16.3. A variável em estudo é:
(I) Habilitações académicas dos trabalhadores da empresa Fonte Nova e é uma variável qualitativa ordinal.
(II) Duração média das baterias para telemóvel de uma certa marca, sendo classificada como variável quantitativa contínua.
(III) Classificação obtida na prova final de Matemática pelos alunos do 9.º ano de uma turma, considerando-se uma variável quantitativa discreta.
(IV) Candidato a Presidente da República que é uma variável qualitativa nominal.
(V) Quantidade de açúcar dos refrigerantes vendidos por uma certa empresa, sendo uma variável quantitativa contínua.

17.
17.1. Neste estudo, optou-se por fazer um censo, pois todos os alunos da escola foram inquiridos.

17.2. A percentagem de alunos que manifestaram preferência por basquetebol é:

$$Percentagem\ basquetebol = 100\% - (18\% + 10\% + 25\% + 12\% + 15\%) = 100\% - 80\% = 20\ \%$$

Concluímos que 20% dos alunos afirmaram preferir basquetebol.

17.3. Temos que o número de alunos que preferem futebol são:

$$N.º\ de\ alunos\ futebol = 200 \times 18\% = 200 \times 0{,}18 = 36$$

Assim, para selecionar 18 alunos para uma amostra aleatória simples, pode-se, por exemplo, numerar os 36 alunos que preferem futebol. De seguida, usando a calculadora, podemos gerar 18 números aleatoriamente de 1 a 36, que corresponderão aos números atribuídos aos 36 alunos. Desta forma, geramos a nossa amostra de 18 alunos.

18.
18.1. A população é o conjunto dos clientes do Mercado Municipal de Trancoso.
18.2. A amostra é o conjunto dos 40 clientes selecionados aleatoriamente.
18.3. A variável em estudo é valor do consumo, em euros, de cada um dos clientes. Trata-se de uma variável quantitativa contínua.
18.4. A percentagem de clientes que gastam menos de 30 euros é:

$$Percentagem = \frac{12+18}{40} \times 100 = \frac{30}{40} \times 100 = 0{,}75 \times 100 = 75\%$$

Concluímos que 75% dos clientes do mercado municipal de Trancoso gastam menos de 30 euros.

18.5. Para determinar graficamente a mediana e os quartis deste conjunto de dados, iremos recorrer à tabela de frequências seguinte.

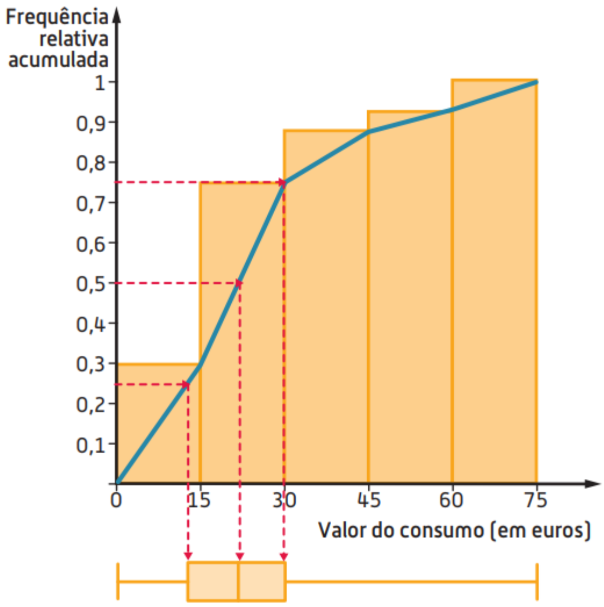
Valor de consumo gasto por 40 clientes do mercado municipal de Trancoso

Valor do consumo (em euros)	Frequência absoluta	Frequência relativa (%)	Frequência relativa acumulada (%)
$[0{,}15[$	12	$\frac{12}{40} = 30\%$	30
$[15{,}30[$	18	$\frac{18}{40} = 45\%$	$30 + 45 = 75$
$[30{,}45[$	5	$\frac{5}{40} = 12{,}5\%$	$75 + 12{,}5 = 87{,}5$
$[45{,}60[$	2	$\frac{2}{40} = 5\%$	$87{,}5 + 5 = 92{,}5$
$[60{,}75[$	3	$\frac{3}{40} = 7{,}5\%$	$92{,}5 + 7{,}5 = 100$

Recorrendo ao histograma e ao polígono de frequências relativas acumuladas, conseguimos determinar graficamente a mediana e os quartis, bem como construir o diagrama de extremos e quartis.

Fonte: Questionário Mercado Municipal de Trancoso

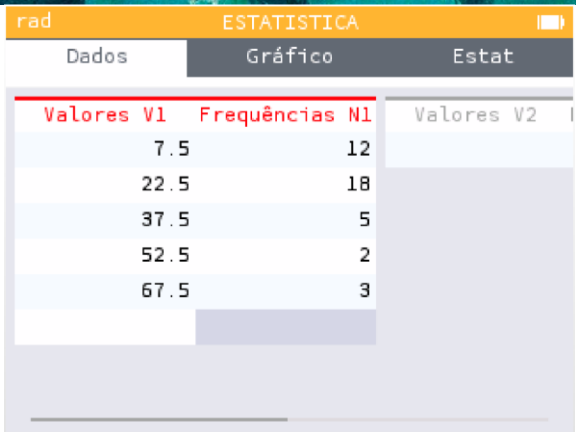
Histograma, polígono de frequências relativas acumuladas e diagrama de extremos e quartis do valor de consumo gasto por 40 clientes do mercado municipal de Trancoso



18.6. Para determinar um valor aproximado da média e do desvio padrão, iremos recorrer às marcas das classes.

Valor do consumo (em euros)	Frequência absoluta	Frequência relativa (%)
$[0,15[$	12	$\frac{0+15}{2}=7{,}5$
$[15,30[$	18	$\frac{15+30}{2}=22{,}5$
$[30,45[$	5	$\frac{30+45}{2}=37{,}5$
$[45,60[$	2	$\frac{45+60}{2}=52{,}5$
$[60,75[$	3	$\frac{60+75}{2}=67{,}5$

Recorrendo às potencialidades de uma calculadora gráfica (NumWorks), começamos por selecionar o menu *Estatística*. De seguida, introduzimos na lista *Valores V1* as marcas das classes e na lista *Frequência N1* as frequências absolutas associadas.

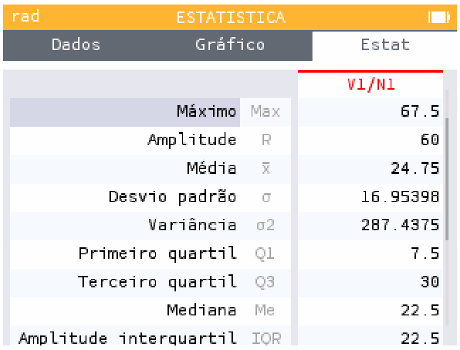


rad ESTATISTICA

Dados | Gráfico | Estat

Valores V1	Frequências N1	Valores V2
7.5	12	
22.5	18	
37.5	5	
52.5	2	
67.5	3	

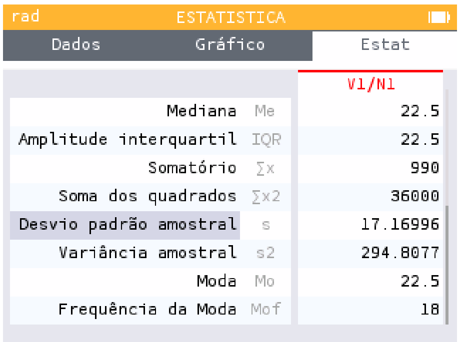
Selecionamos a seguir a opção *Estat*, ao que obtemos os valores pretendidos.



rad ESTATISTICA

Dados | Gráfico | Estat

		V1/N1
Máximo	Max	67.5
Amplitude	R	60
Média	x̄	24.75
Desvio padrão	σ	16.95398
Variância	σ2	287.4375
Primeiro quartil	Q1	7.5
Terceiro quartil	Q3	30
Mediana	Me	22.5
Amplitude interquartil	IQR	22.5



rad ESTATISTICA

Dados | Gráfico | Estat

		V1/N1
Mediana	Me	22.5
Amplitude interquartil	IQR	22.5
Somatório	∑x	990
Soma dos quadrados	∑x2	36000
Desvio padrão amostral	s	17.16996
Variância amostral	s2	294.8077
Moda	Mo	22.5
Frequência da Moda	Mof	18

$$\bar{x} = 24{,}75 \qquad s \approx 17{,}2$$

O valor do consumo médio é de 24,75 euros e o desvio padrão é, aproximadamente, 17,2 euros.

19.

19.1. Podemos por excluir a opção (A), pois apenas estão representados 12 elementos no total.

A amplitude é igual à diferença entre o máximo e o mínimo. Assim, para cada uma das opções (B), (C) e (D), a amplitude é a seguinte.

Opção (A) $\quad Amplitude = 30 - 18 = 12$

Opção (B) $\quad Amplitude = 31 - 12 = 19$

Opção (D) $\quad Amplitude = 31 - 12 = 19$

Como apenas as opções (B) e (D) tem um valor de amplitude igual a 19, podemos excluir a opção (A).

A média de cada conjunto de dados das opções (B) e (D) é:

$$\bar{x}_B = \frac{12 \times 2 + 13 \times 2 + 15 \times 3 + 20 + 21 + 22 + 30 \times 2 + 31 \times 2}{14} = 20$$

$$\bar{x}_D = \frac{12 \times 3 + 13 \times 2 + 15 \times 2 + 17 + 18 + 19 + 21 \times 2 + 22 + 31}{14} \approx 17{,}2$$

A média do conjunto de dados é 19 na opção (B).

19.2. Seja x o número de elementos do grupo B. Temos então que:

- $N.º\ total\ de\ pessoas = 14 + x$
- $Soma\ das\ idades\ de\ todos = 20 \times 14 + 18 \times x = 280 + 18x$

Como a média do grupo constituído pelos $14 + x$ elementos é 18,7, podemos escrever a seguinte equação.

$$\bar{x} = 18{,}7 \Leftrightarrow \frac{280 + 18x}{14 + x} = 18{,}7 \Leftrightarrow$$
$$\Leftrightarrow 280 + 18x = 18{,}7 \times (14 + x) \Leftrightarrow$$
$$\Leftrightarrow 18x - 18{,}7x = 261{,}8 - 280 \Leftrightarrow$$
$$\Leftrightarrow x = \frac{18{,}2}{0{,}7} \Leftrightarrow x = 26$$

O grupo B é constituído por 26 pessoas.

20.

20.1. Fazendo um levantamento dos dados registados, podemos representar o diagrama de caule-e-folhas seguinte.

Tempo de estudo a Matemática A, em minutos

0	0 0
1	0 0 0
2	0 5
3	0 0 5
4	5 5 5
5	0 0 5
6	0 0
7	0 0 5
8	0 5
9	0 5

0|0 significa 0 minutos

Fonte: Própria

20.2. Os percentis 25, 50 e 75 correspondem aos valores dos 1.º e 3.º quartis e da mediana.
Comecemos por construir a tabela de frequências seguinte.

Tempo de estudo a Matemática A, em minutos, de uma turma do 10.º ano

Tempo (em minutos)	Frequência absoluta	Frequência relativa (%)	Frequência relativa acumulada (%)
0	2	8	8
10	3	12	20
20	1	4	24
25	1	4	28
30	2	8	36
35	1	4	40
45	3	12	52
50	2	8	60

Tempo (em minutos)	Frequência absoluta	Frequência relativa (%)	Frequência relativa acumulada (%)
55	1	4	64
60	2	8	72
70	2	8	80
75	1	4	84
80	1	4	88
85	1	4	92
90	1	4	96
95	1	4	100

Tem-se que:

$$P_{25} = Q_1 = 25 \qquad P_{50} = M_e = 45 \qquad P_{75} = 70$$

Concluímos que os percentis 25, 50 e 75 deste conjunto de dados são 25, 45 e 70 minutos, respetivamente.

20.3. Recorrendo à tabela construída anteriormente, o tempo 60 minutos pertence aos $P_{65}, P_{66}P_{67}, P_{68}, P_{69}, P_{70}, P_{71}$ e P_{72}.

21.

21.1. A variável é quantitativa discreta.

21.2. Fazendo um levantamento dos dados recolhidos, construímos o diagrama de caule-e-folhas seguinte.

N.º de vezes de utilização da cafetaria da escola do Martim, durante o mês

0	0 0 1 2 2 3 4 5 5 6 8 9
1	0 0 0 1 2 2 2 3 6 6 8
2	0 0 2 2 2 2 2 4 4 5
3	5
4	0 0 2 4 4
5	2
6	0 4

0|0 significa 0 minutos

Fonte: Escola do Martim

21.3. A tabela é completada da forma seguinte.

Número de vezes que 42 alunos frequentaram a cafetaria da escola

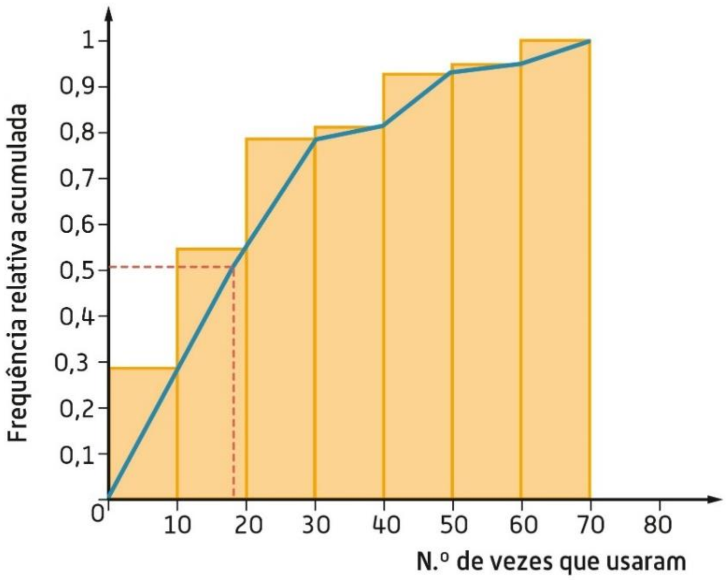
Número de vezes	Frequência absoluta	Frequência relativa
[0,10[	12	$\frac{12}{42}=\frac{2}{7}$
[10,20[	11	$\frac{11}{42}$
[20,30[	10	$\frac{10}{42}=\frac{5}{21}$
[30,40[	1	$\frac{1}{42}$
[40,50[	5	$\frac{5}{42}$
[50,60[	1	$\frac{1}{42}$
[60,70[	2	$\frac{2}{42}=\frac{1}{21}$

21.4. Para determinar um valor aproximado do percentil do número de vezes que o Martim foi à cafetaria da escola, podemos de calcular as frequências relativas acumuladas. Neste caso, iremos representar essas frequências em forma de percentagem, arredondada às centésimas.

Classe	Frequência absoluta	Frequência relativa	Frequência relativa acumulada (%)
[0,10[	12	$\frac{12}{42}=\frac{2}{7}$	28,57%
[10,20[	11	$\frac{11}{42}$	54,76%
[20,30[	10	$\frac{10}{42}=\frac{5}{21}$	78,57%
[30,40[	1	$\frac{1}{42}$	80,95%
[40,50[	5	$\frac{5}{42}$	92,86%
[50,60[	1	$\frac{1}{42}$	95,24%
[60,70[	2	$\frac{2}{42}=\frac{1}{21}$	100%

De seguida, podemos construir o histograma de frequências relativas acumuladas, bem como o respetivo polígono de frequências.

Histograma e polígono de frequências relativas acumuladas do n.º de vezes que os alunos utilizaram a cafetaria da escola



Fonte: Escola do Martim

O percentil correspondente ao número de vezes que o Martim usou a cafetaria é P_{51}, aproximadamente.

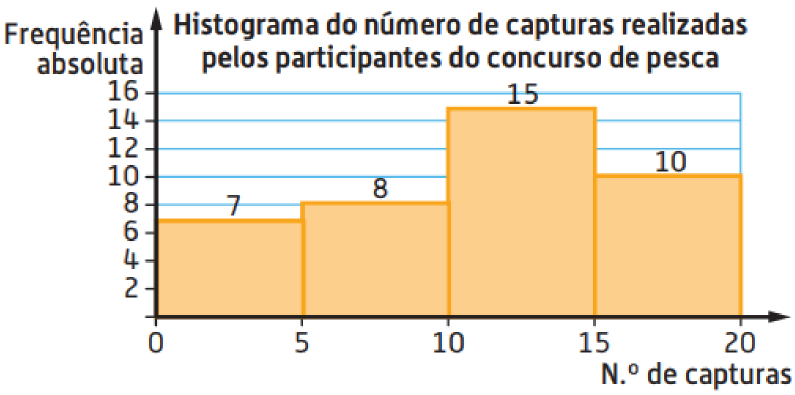
22.

22.1. O número de pescadores participantes neste concurso foi:

$$N.º\ de\ participantes = 7 + 8 + 15 + 10 = 40$$

Participaram no concurso 40 pescadores.

22.2.



Fonte: Concurso

22.3. O número de capturas do Joaquim foi considerado na classe $[15, 20[$.

22.4. Para determinar a média, consideramos as marcas de cada uma das classes.

Número de capturas de cada participante no concurso de pesca

Número de capturas	Frequência absoluta	Marca da classe
$[0,5[$	7	$\frac{0+5}{2}=2{,}5$
$[5,10[$	8	$\frac{5+10}{2}=7{,}5$
$[10,15[$	15	$\frac{10+15}{2}=12{,}5$
$[15,20[$	10	$\frac{15+20}{2}=17{,}5$
Total	40	

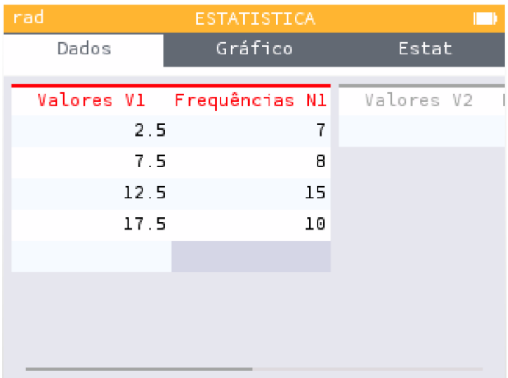
Um valor aproximado para a média do número de capturas é:

$$\bar{x}=\frac{2{,}5\times 7+7{,}5\times 8+12{,}5\times 15+17{,}5\times 10}{40}=11$$

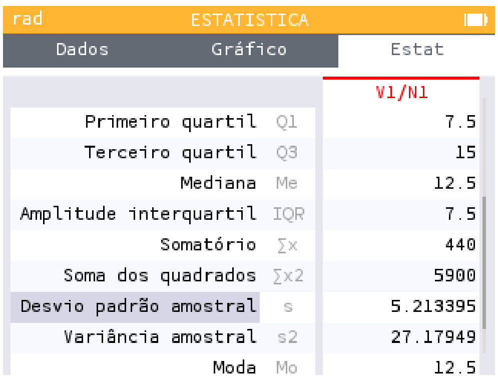
O número de capturas médio é 11.

22.5. Usando as potencialidades de uma calculadora gráfica (*NumWorks*), começamos por selecionar o menu Estatística.

De seguida, introduzimos os dados da tabela nas listas *Valores V1* e *Frequências N1*.



Selecionando a opção *Estat*, obtemos assim o valor do desvio padrão seguinte.



$s\approx 5{,}21$

O valor aproximado do desvio padrão é 5,21 capturas.

22.6. A classe modal é [10, 15[.

Para determinar a classe mediana, iremos recorrer à tabela de frequências seguinte.

Número de capturas de cada participante no concurso de pesca

Número de capturas	Frequência absoluta	Frequência relativa (%)	Frequência relativa acumulada (%)
[0, 5[	7	$\frac{7}{40} \times 100 = 17{,}5$	17,5%
[5, 10[	8	$\frac{8}{40} \times 100 = 20$	37,5%
[10, 15[	15	$\frac{15}{40} \times 100 = 37{,}5$	75%
[15, 20[	10	$\frac{10}{40} \times 100 = 25$	100%

Podemos então concluir que a classe mediana é [10, 15[.

23. Como a média das primeiras quatro avaliações é 17,2, então a soma das primeiras quatro classificações é:

$$Soma = 4 \times 17{,}2 = 68{,}8$$

Assim, a média das classificações obtidas nestas cinco avaliações é:

$$\bar{x} = \frac{68{,}8 + 18}{5} = \frac{86{,}8}{5} = 17{,}36$$

O Gilberto teve uma classificação média de 17,36 valores na disciplina de Filosofia.

24.

24.1. A amplitude é igual à diferença entre a classificação máxima e a classificação mínima obtidas pelo Manuel.

Assim:

$$mín. = 2 \qquad\qquad máx. = 5$$

$$Amplitude = 5 - 2 = 3$$

A amplitude das classificações do Manuel é igual a 3.

24.2. Para determinar a amplitude interquartil, começamos por determinar os 1.º e 3.º quartis.
Para isso, vamos começar por ordenar as classificações, por ordem crescente.

2	2	3	3	4	5	5	5
	$Q_1 = \frac{2+3}{2} = 2{,}5$				$Q_3 = \frac{5+5}{2} = 5$		

A amplitude interquartil do conjunto de classificações do Manuel é:

$$Amplitude\ interquatil = Q_3 - Q_1 = 5 - 2{,}5 = 2{,}5$$

A amplitude interquartil deste conjunto de dados é 2,5.

25. Comecemos por ordenar as frequências cardíacas, por ordem crescente.

66	67	67	68	70	70	71	71	72	72	72	75	75	77	77	79	79	80	80	85
				valores centrais da 1.ª metade										valores centrais da 2.ª metade					

Podemos então determinar que:

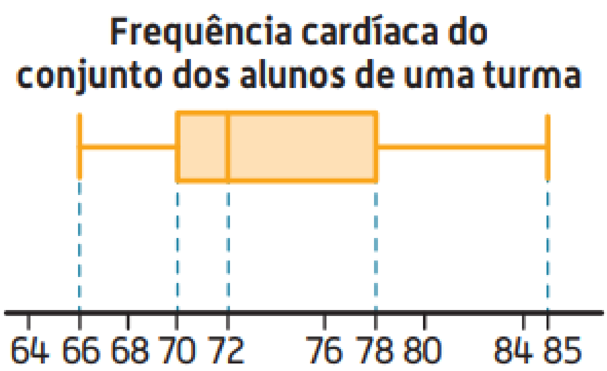
$$min. = 66 \qquad máx. = 85$$

$$Q_1 = \frac{70+70}{2} = 70 \qquad M_e = \frac{72+72}{2} = 72 \qquad Q_3 = \frac{77+79}{2} = 78$$

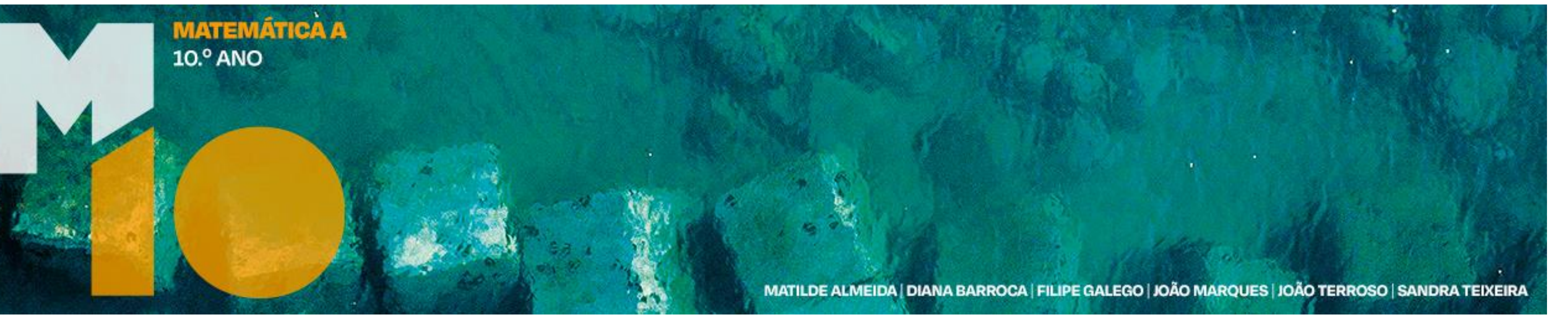
Os valores da amplitude e amplitude interquartil são:

- $Amplitude = máx. - mín. = 85 - 66 = 19$
- $Amplitude\ interquartil = Q_3 - Q_1 = 78 - 70 = 8$

Representamos o diagrama de extremos e quartis do conjunto de frequências cardíacas recolhidas da seguinte forma.



Fonte: Registos dos alunos da turma



26. No caso de o Conselho de Administração aumentar o salário de cada trabalhador em 250 €, a média aumentará 250 €, mas o desvio padrão manter-se-á. Assim, a média passará a 1900€ e o desvio padrão será de 231,5€.

Se os vencimentos de cada um dos trabalhadores aumentarem 25 % (ou seja, se o salário de cada trabalhador passar a ser 1,25 vezes superior ao salário atual), então a média passará a ser $1650€ \times 1{,}25 = 2062{,}5€$ e o desvio padrão $231{,}5€ \times 1{,}25 \approx 289{,}38€$.

27.

27.1. As médias de cada uma das turmas são:

$$\bar{x}_A = \frac{1 \times 2 + 2 \times 2 + 3 \times 8 + 4 \times 4 + 5 \times 2}{20} = 3$$

$$\bar{x}_A = \frac{1 \times 6 + 2 \times 3 + 3 \times 2 + 4 \times 3 + 5 \times 6}{20} = 3$$

As médias das classificações das turmas A e B coincidem e são iguais a 3.

27.2.

Recorrendo às potencialidades de uma calculadora gráfica (*NumWorks*), selecionamos o menu *Estatística* e introduzimos nas listas os resultados das turmas 9.º A e 9.º B, bem como das respetivas frequências absolutas.

rad ESTATISTICA — Dados | Gráfico | Estat

Valores V1	Frequências N1	Valores V2
1	2	1
2	4	2
3	8	3
4	4	4
5	2	5

rad ESTATISTICA — Dados | Gráfico | Estat

Valores V2	Frequências N2	Valores V3	Fr
1	6		
2	3		
3	2		
4	3		
5	6		

Depois de introduzidos os dados, selecionamos *Estat*, pelo que a máquina responde com os valores seguintes do desvio padrão.

rad ESTATISTICA — Dados | Gráfico | Estat

		V1/N1	V2/N2
Mediana	Me	3	3
rquartil	IQR	2	4
omatório	∑x	60	60
uadrados	∑x2	204	234
amostral	s	1.123903	1.685854
amostral	s2	1.263158	2.842105
Moda 1	Mo1	3	1
Moda 2	Mo2	-	5
da Moda	Mof	8	6

$s_A \approx 1{,}124$ $\qquad s_B \approx 1{,}686$

A turma B apresenta um maior valor do desvio padrão, ou seja, houve maior dispersão nas classificações obtidas pela turma B do que pela turma A.

28.

28.1. A equipa com maior dispersão das alturas é a equipa B.

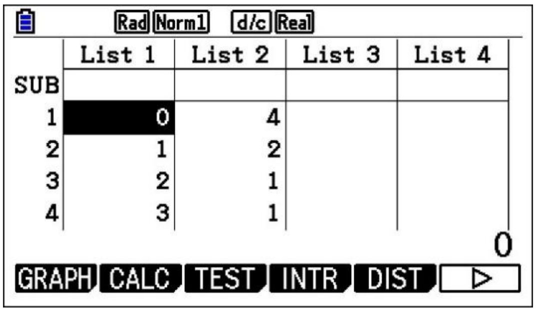
28.2. Observando os histogramas e analisando o conjunto de médias e desvios padrão, podemos concluir que:

I – Equipa D II – Equipa B III – Equipa C IV – Equipa A

29.

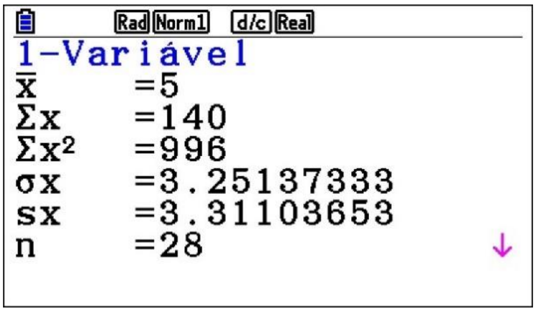
29.1. Usando as potencialidades da calculadora gráfica (*Casio fx-CG50*), começamos por selecionar a opção *Estatística*.

De seguida, na *Lista 1* introduzimos os valores da variável *número de faltas da turma C* e na *Lista 2* introduzimos o *número de alunos* associado.



De seguida, selecionamos a tecla F2 CALC e, logo a seguir, F6 SET. Verificamos as entradas 1 Var XLIST : List1 e 1Var Freq: List2.

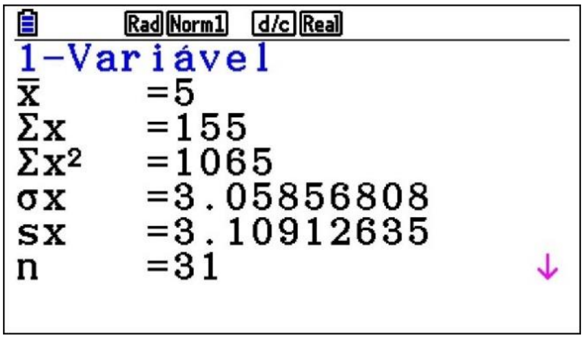
Depois de realizados estes procedimentos, selecionamos EXIT, seguido de F1 1-VAR.



A máquina devolve os seguintes valores da média e desvio padrão do conjunto de faltas da turma C.

$\bar{x}_C = 5$ $s \approx 3{,}31$

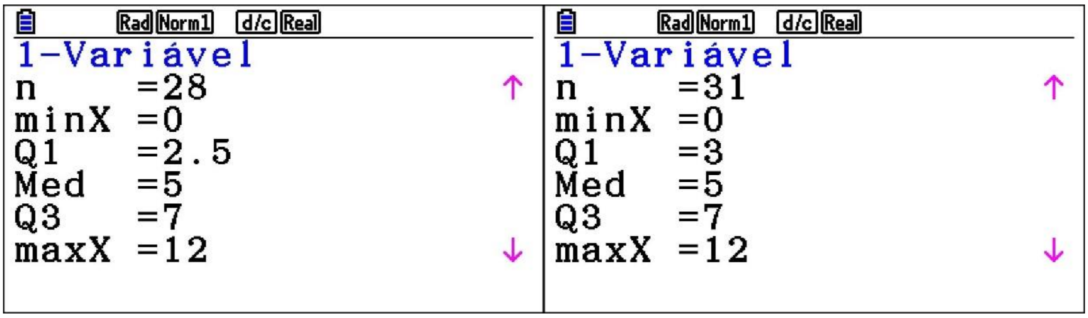
Usando o mesmo procedimento para a turma D, obtemos os seguintes valores da média e do desvio padrão do conjunto de faltas da turma D.



$\bar{x}_D = 5$ $\qquad\qquad$ $s \approx 3{,}11$

Como os valores do número médio de faltas das duas turmas são iguais e os respetivos desvios padrão muito próximos, não se conseguem indicar diferenças entre as duas turmas.

29.2. Usando as potencialidades da calculadora gráfica, obtemos os valores dos extremos e quartis para cada uma das turmas.

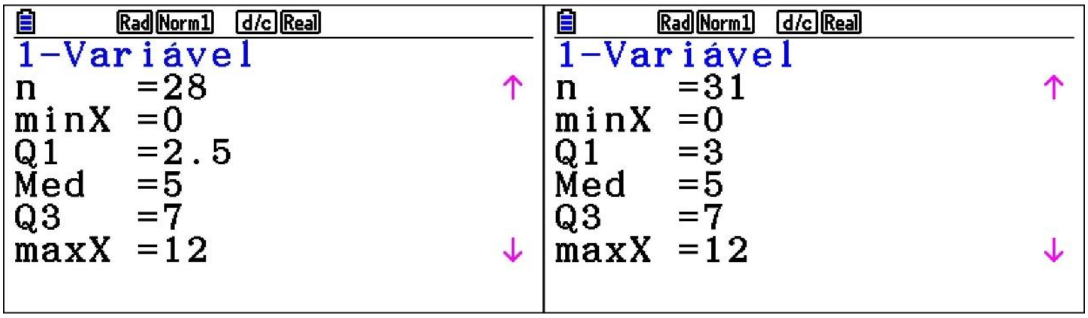


Turma C

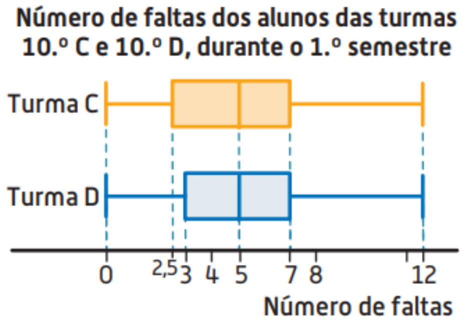
$mín. = 0$ $\qquad$ $Q_1 = 2{,}5$ $\qquad$ $M_e = 5\, Q_3 = 7\ máx. = 12$

Turma D

$mín. = 0$ $\qquad$ $Q_1 = 3$ $\qquad$ $M_e = 5\, Q_3 = 6\ máx. = 12$



Com estes valores, representamos os diagramas de extremos e quartis paralelos seguintes.



Fonte: Registos de faltas dos alunos do 10.º C e 10.º D

Na turma C, os dados centrais estão mais dispersos e há uma maior dispersão entre o 1.º quartil e a mediana do que entre a mediana e o terceiro quartil. Na turma D, os dados centrais estão mais concentrados e a concentração dos dados entre o 1.º quartil e a mediana ser igual à concentração entre a mediana e o terceiro quartil.

30.
30.1. Opção (B).

30.2. Observando os diagramas de extremos e quartis, concluímos que: na empresa A, os 25% dos salários mais altos concentram-se entre os 1600 euros e os 2000 euros; na empresa B, 25% dos salários mais altos estão compreendidos entre 1600 euros e 2200 euros; na empresa C, os 25% dos salários mais altos estão entre 1800 euros e 2200 euros.

31.
31.1. Recorrendo a folha de cálculo e considerando os pares ordenados (x_i, y_i) (em que x_i corresponde aos valores da variável número de anos no grupo e y_i corresponde aos valores da variável ordenado bruto, em euros), podemos construir o diagrama de dispersão e a respetiva reta de regressão da forma seguinte.

Fonte: ValorTempo – Gestão do Tempo e Horas, SA

O coeficiente de correlação é:

- $r \approx ,995$

Trata-se de uma correlação positiva forte, pois o coeficiente de correlação é próximo de 1.

31.2. Tendo em conta que existe uma forte correlação positiva entre as variáveis, podemos afirmar que as pessoas com mais anos no grupo têm tendencialmente ordenados mais elevados.

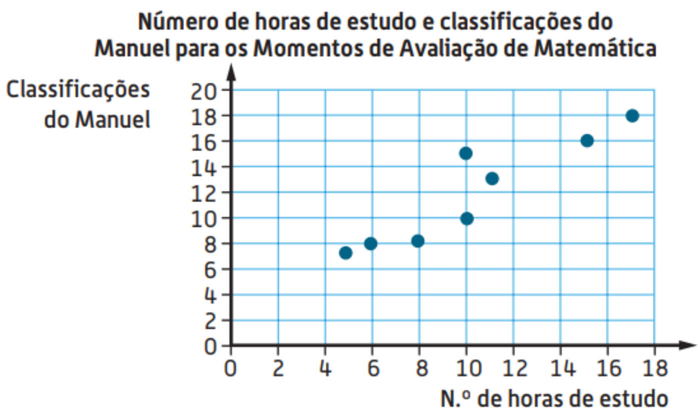
31.3. A variável independente ou explanatória é o número de anos no grupo empresarial.

31.4. A variável dependente ou resposta é o ordenado bruto, em euros.

32.

32.1. Seja o conjunto de pares ordenados (x_i, y_i) tais que x_i corresponde aos valores do variável número de horas de estudo e y_i corresponde aos valores da variável classificação do Matemática do Manuel.

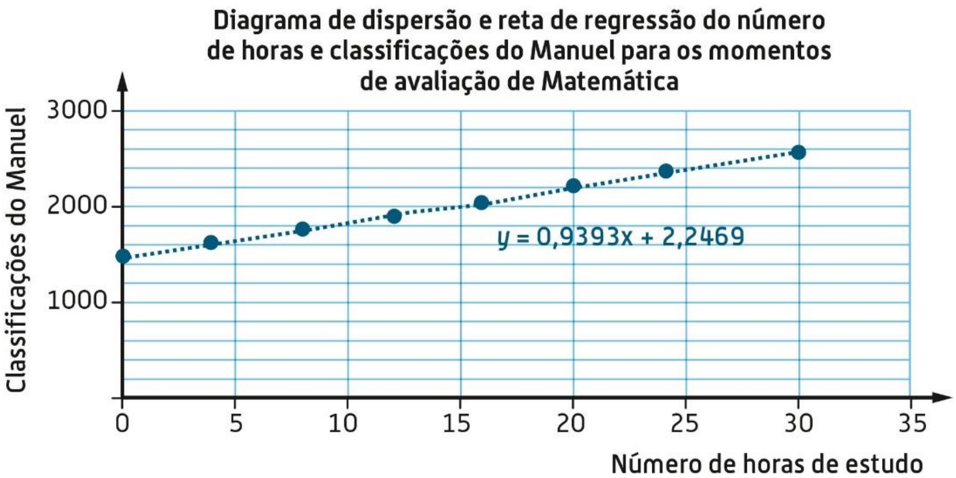
Podemos representar o diagrama de dispersão da forma seguinte.



Fonte: Momentos de avaliação do Manuel

32.2. Seja $y = ax + b$, com a e b constantes reais, a equação da reta de regressão.

Usando as potencialidades de uma folha de cálculo, obtemos a seguinte representação da reta de regressão.



Fonte: Momentos de avaliação do Manuel

Temos que:

$$r \approx 0{,}939 \qquad a \approx 0{,}939 \qquad b \approx 2{,}247$$

A reta de regressão é definida pela equação $y = 0{,}939x + 2{,}247$ e o coeficiente de correlação é, aproximadamente, $r \approx 0{,}939$.

32.3. Por hipótese, a classificação do Manuel é de 19 valores ($y = 19$). Então:

$$19 = 0{,}9393x + 2{,}2469 \Leftrightarrow 0{,}9393x = 19 - 2{,}2469 \Leftrightarrow x \approx 18$$

Podemos estimar que o número de horas de estudo é 18.

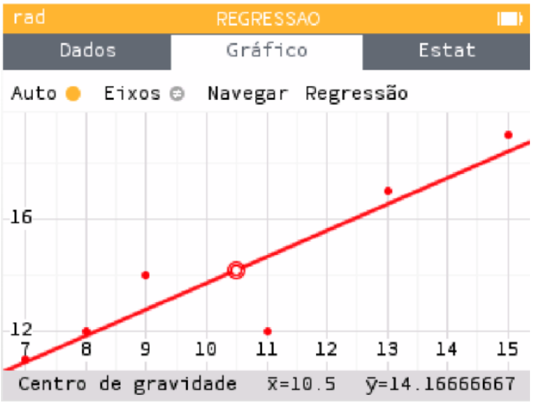
33.

33.1. Vamos utilizar a calculadora gráfica (*NumWorks*) para representar as variáveis x e y num diagrama de dispersão.

Para isso começamos por selecionar no menu principal regressão e copiar os valores da tabela.

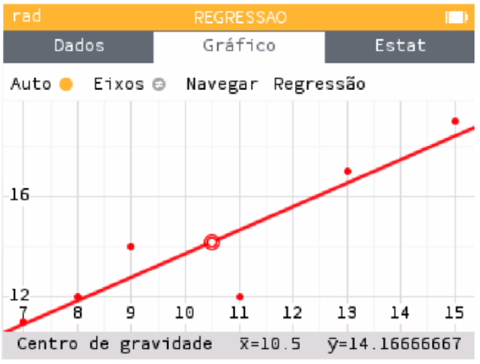
X1	Y1	X2
13	17	
9	14	
15	19	
11	12	
7	11	
8	12	

Obtemos o seguinte diagrama de dispersão.



33.2. Através da máquina de calcular, conseguimos determinar o centro de gravidade. As coordenadas do ponto de gravidade são $(10{,}5; 14{,}2)$.

Representando o par ordenado $(10{,}5; 14{,}2)$ no diagrama de dispersão, temos que:



33.3.

Selecionando o separador *Estatística*, obtermos o valor do coeficiente de correlação.

O coeficiente de correlação é, aproximadamente, $0{,}906$. Assim, como o valor é próximo de 1, podemos afirmar que se trata de uma correlação positiva forte.

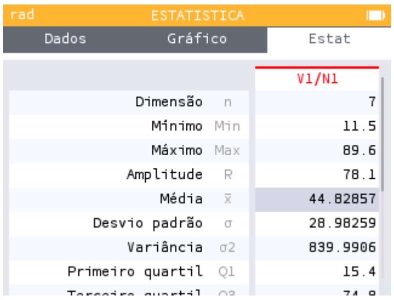
34.

34.1.

Usando as potencialidades da calculadora gráfica (*NumWorks*), começamos por selecionar o menu *Estatística*.

A seguir, inserimos na lista *Valores V1* as taxas de alfabetização de adultos da tabela e mantemos a lista *Frequências N1* predefinida.

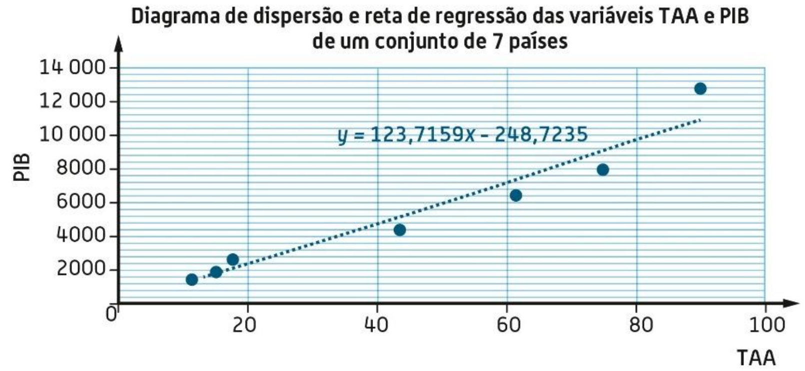
Selecionamos *Estat* e obtemos o valor da média seguinte.



$$\bar{x} \approx 44{,}8$$

A média da taxa de alfabetização de adultos, no conjunto dos setes países, é 44,8%, aproximadamente.

34.2. Usando as potencialidades de uma folha de cálculo, consideremos o conjunto de pares ordenados (x_i, y_i) em que x_i corresponde aos valores da variável TAA e y_i corresponde aos valores da variável PIB.
Com o conjunto de pares ordenados, conseguimos representar o diagrama de dispersão e reta de regressão da forma seguinte.



Fonte: Própria

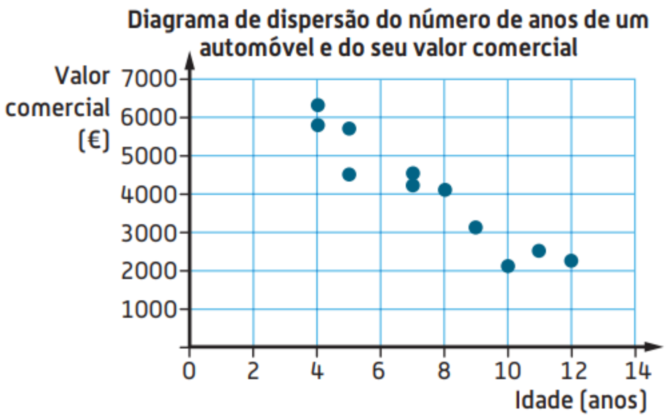
O valor do coeficiente de correlação é:

$$r \approx 0{,}9634$$

A reta de regressão é definida pela equação $y = 123{,}7159x - 248{,}7235$ e o coeficiente de correlação é, aproximadamente, $r \approx 0{,}9634$.

35.

35.1. Considerando os pares ordenados (x_i, y_i) o conjunto de pares ordenados, com x_i o conjunto de valores da variável idade e y_i o conjunto de valores da variável valor comercial (em euros), o diagrama de dispersão pode ser representado da seguinte forma.



Fonte: Própria

35.2. Usando as potencialidades de uma folha de cálculo, temos que o coeficiente de correlação é:

$$r \approx 0{,}955$$

Podemos concluir que existe uma correlação linear negativa forte entre as variáveis idade (em anos) e o valor comercial (em euros) no conjunto de automóveis.

35.3. Usando uma folha de cálculo, temos que:

$$a \approx -501{,}97 \qquad\qquad b \approx 7833{,}8$$

A equação da reta de regressão é $y = -501{,}97x + 7833{,}8$.

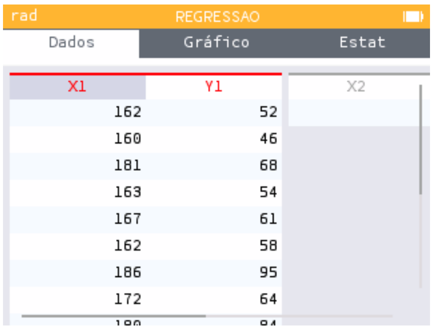
35.4. A variável independente é a idade (em anos) e a variável dependente é o valor comercial (em euros).

36.

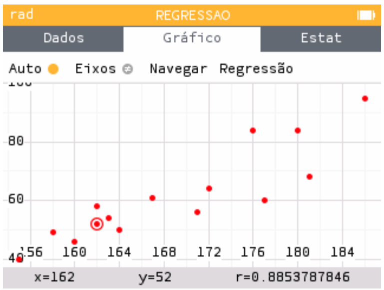
36.1.

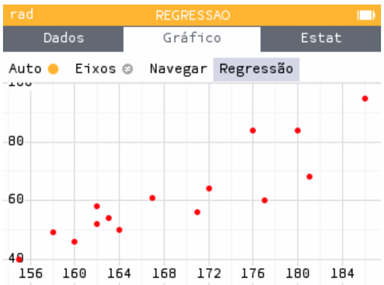
Usando as potencialidades da calculadora gráfica (*NumWorks*), começamos por selecionar o menu *Regressão*.

De seguida, inserimos os dados da tabela nas listas *X1* e *Y1*.



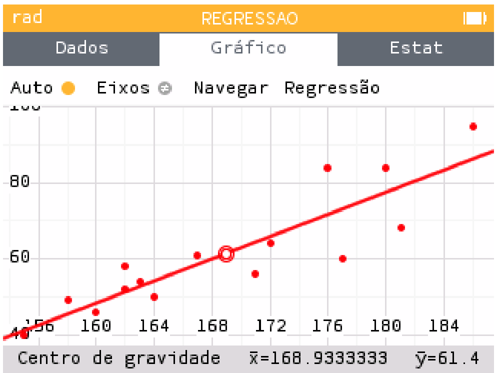
Depois de introduzidas listas, selecionamos a opção *Gráfico*, *Regressão* e, por último, *Linear*.





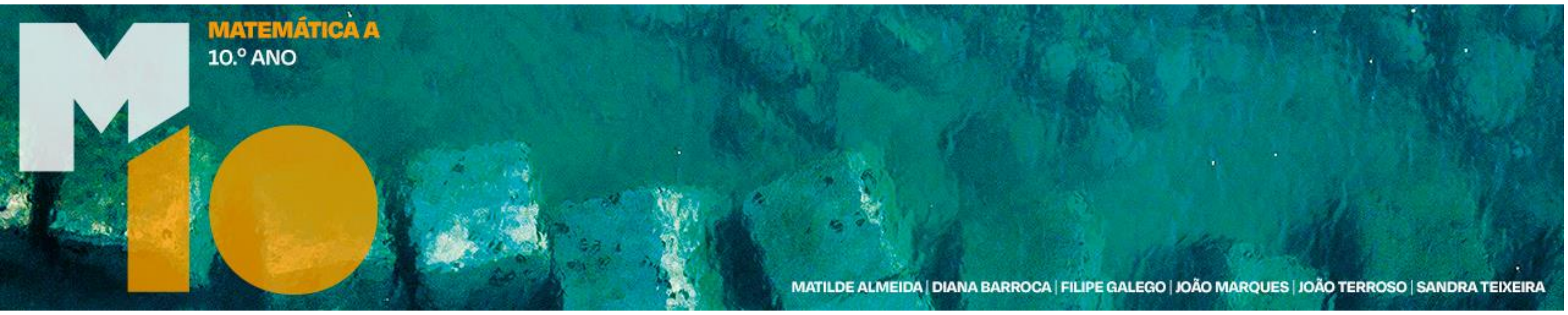
Surge então na janela a representação gráfica da reta de regressão.

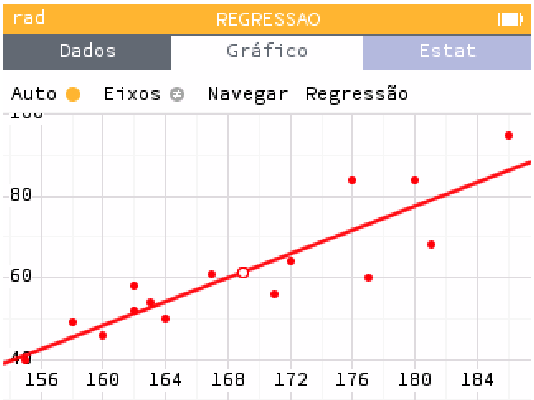
Reta de regressão das variáveis "altura (em cm)" e "massa (em kg)" do conjunto de 15 pessoas



Fonte: Própria

Para obter o coeficiente de correlação e parâmetros m e b da equação $y = mx + b$ da reta de regressão, selecionamos *Estat*, e percorremos a lista de estatísticas da calculadora.





Obtemos então os valores seguintes.

$$r \approx 0{,}89 \qquad\qquad a \approx 1{,}46 \qquad\qquad b \approx -184{,}92$$

O coeficiente de correlação é, aproximadamente, 0,89 e a equação da reta de regressão é

$$y = 1{,}46x - 184{,}92.$$

36.2. Se a altura de uma pessoa é 180 cm ($x = 180$), então a massa estimada é:

$$y = \ 1{,}46 \times 180 - 184{,}92 = 77{,}88\ kg$$

No caso da altura de uma pessoa ser de 1,80 m, estima-se que a sua massa seja 77,88 kg.

36.3. No caso da massa de uma pessoa ser 70 kg ($y = 70$), a altura previsível corresponde à solução da equação $1{,}46x - 184{,}92 = 70$.

Assim:

$$1{,}46x \ - \ 184{,}92 = 70 \Leftrightarrow 1{,}46x = 70 - 184{,}92 \Leftrightarrow x = \frac{70 + 184{,}92}{1{,}458} \Leftrightarrow x \approx 175$$

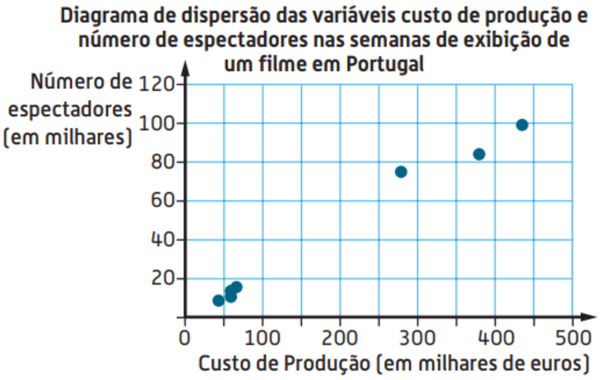
Estima-se que uma pessoa com 70 kg de massa corporal meça 174,8 cm.

37.

37.1.

Considera o conjunto de pares ordenados (x_i, y_i), onde x_i corresponde aos valores da variável custo de produção e y_i corresponde aos valores da variável número de espetadores.

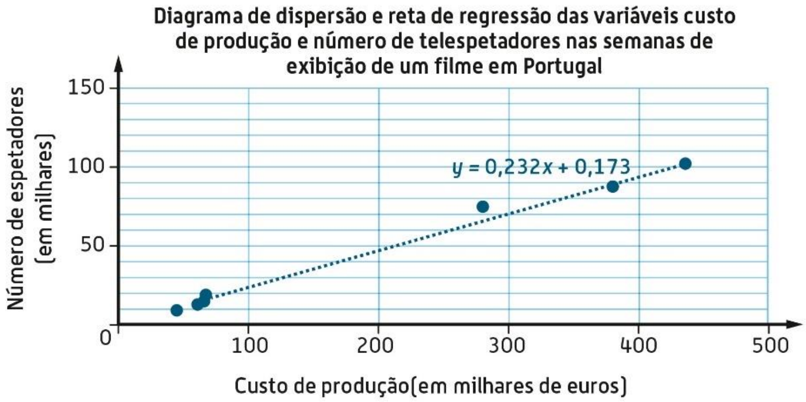
A representação gráfica do diagrama de dispersão do conjunto de pares ordenados é ao lado.



Fonte: Própria

37.2. Considera a equação da reta de regressão $y = ax + b$, com a e b constantes reais.

Usando as potencialidades de uma folha de cálculo, conseguimos obter os valores dos parâmetros a e b da reta. Assim:



Fonte: Própria

$$a \approx 0{,}232 \qquad\qquad b \approx 0{,}173$$

A equação da reta de regressão é definida por $y = 0{,}232x + 0{,}173$.

37.3. O número estimado de espectadores de um filme com um custo de produção de 75 mil euros ($x = 75$) é:

$$y \; = \; 0{,}232 \times 75 \; + \; 0{,}173 = 17{,}573$$

O número esperado de espectadores para assistir a um filme com um custo de produção de 75 milhares de euros é, aproximadamente ,17 573.

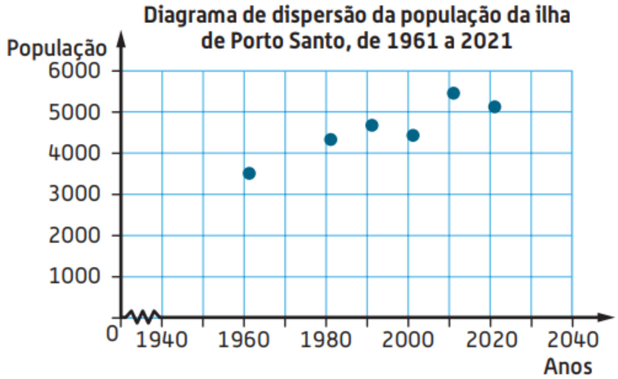
37.4. Por hipótese, o número de espectadores é 52,5 milhares, então o custo estimado da produção é a solução da equação $52{,}5 = \; 0{,}232x \; + \; 0{,}173$. Assim:

$$0{,}232x \; + \; 0{,}173 = 52{,}5 \Leftrightarrow 0{,}232x = 52{,}5 - 0{,}173 \Leftrightarrow x = \frac{52{,}5 - 0{,}173}{0{,}232} \Leftrightarrow x \approx 225{,}54741$$

Estima-se que um filme com 52,5 milhares de telespectadores a assistir terá como custo de produção 225 547,41 euros.

38.

38.1. Seja o conjunto de pares ordenados (x_i, y_i), em que x_i corresponde aos valores da variável ano e y_i corresponde aos valores da variável população.



Fonte: Censos

O conjunto de pares ordenados podem ser representados no diagrama de dispersão seguinte.

38.2. Usando as potencialidades de uma folha de cálculo, podemos determinar os valores de m e b. Assim:

$$m \approx 29{,}0 \qquad\qquad b \approx -53\,205{,}9$$

A equação da reta de regressão é $y = 29x - 53\,205{,}9$.

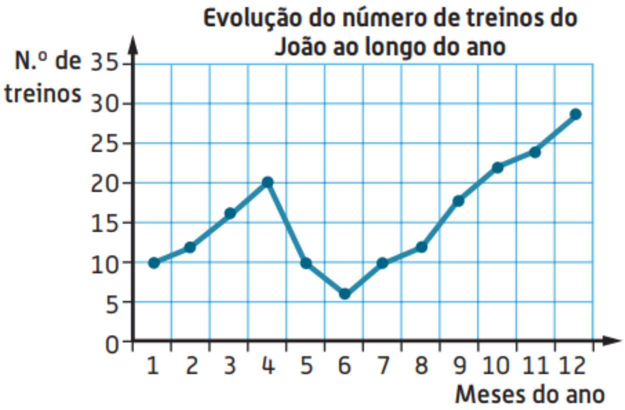
38.3. Tendo em consideração que $x = 2031$, podemos estimar que:

$$y = 29 \times 2031 - 53\,205{,}9 \approx 5693$$

Em 2031, estima-se que na ilha de Porto Santo haja um total de 5693 habitantes

39.

39.1. Considerando x a variável mês e y a variável número de treinos realizados, conseguimos representar o gráfico de linhas seguinte.



Fonte: Registos do João

39.2. A média do número de treinos mensais é:

$$\bar{x} = \frac{10 + 12 + 16 + 20 + 10 + 6 + 10 + 12 + 18 + 22 + 24 + 29}{12} = 15{,}75$$

O número médio de treinos realizados mensalmente pelo João foi de 15,75.

39.3. Começamos por determinar a variância da forma seguinte.

$$s^2 = \frac{(10 - 15{,}75)^2 + (12 - 15{,}75)^2 + \cdots + (24 - 15{,}75)^2 + (29 - 15{,}75)^2}{12 - 1} \approx 48{,}023$$

O desvio padrão do número de treinos é:

$$s = \sqrt{48{,}023} \approx 6{,}93$$

O desvio padrão é, aproximadamente, 6,93.

40.

40.1. Entre o 6.º e o 7.º mês, a taxa de utilização da cantina diminui de 12,7 % para 9,4 %. Ou seja, sabemos que a diminuição da taxa de utilização foi de 3,3%.

Assim:

$$\frac{12{,}7\%}{100\%}=\frac{3{,}3}{x}\Leftrightarrow x=\frac{3{,}3\times 100}{12{,}7}\Leftrightarrow x\approx 26\ \%$$

Opção (A)

40.2. Comecemos por ordenar as taxas de utilização da cantina no total dos 12 meses.

b	9,4	11,7	12,7	13,9	14,5	a	15,5	15,9	15,9	16,2	16,5

Assim, sabemos que a mediana se obtém fazendo a média dos dois valores centrais.

b	9,4	11,7	12,7	13,9	14,5	a	15,5	15,9	15,9	16,2	16,5
					$M_e=\frac{14{,}5+a}{2}$						

Por hipótese, sabemos que a mediana é igual a 14,9. Assim:

$$M_e=14{,}9\Leftrightarrow\frac{14{,}5+a}{2}=14{,}9\Leftrightarrow 14{,}5+a=2\times 14{,}9\Leftrightarrow a=29{,}8-14{,}5\Leftrightarrow a=15{,}3$$

O valor de a é 15,3%.

Avaliar - Página 178 e 179

1.

1.1. A empresa optou por fazer uma sondagem, pois seria demorado e oneroso inquirir toda a população da região norte de Portugal.

1.2. A população em estudo é o conjunto de habitantes da região norte de Portugal.

1.3. A amostra tem dimensão igual a 1560 pessoas.

1.4. Podemos afirmar que se usou uma amostra aleatória simples.

2.

2.1. Com os registos das classificações da primeira avaliação de Matemática da turma da Francisca, conseguimos contruir o diagrama de caule-e-folhas seguinte.

Classificações (em percentagem) da primeira avaliação de Matemática na turma da Francisca

1	2
2	2 7
3	2 8
4	5
5	
6	0 6
7	7 7
8	2 4 5 8
9	0 8

1|2 significa 12

Fonte: Registos do Professor

2.2. A classificação média da turma na primeira prova de Matemática é:

$$\bar{x} = \frac{12 + 22 + 27 + 32 + 38 + 45 + 60 + 66 + 77 \times 2 + 82 + 84 + 85 + 88 + 90 + 98}{16} \approx 61{,}44$$

Para determinar a mediana, iremos ter de ordenar os dados, de forma crescente.

12	22	27	32	38	45	60	66	77	77	82	84	85	88	90	98

$$M_e = \frac{66 + 77}{2} = 71{,}5$$

Para determinar o desvio padrão, teremos de calcular a variância:

$$s^2 = \frac{(12 - 49{,}15)^2 + (22 - 49{,}15)^2 + \cdots + (90 - 49{,}15)^2 + (98 - 49{,}15)^2}{16 - 1} = 942{,}644$$

O desvio padrão é:

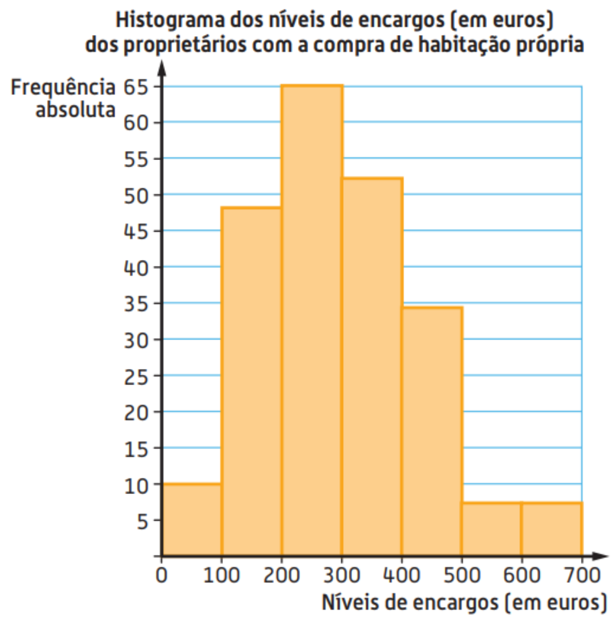
$$s = \sqrt{942{,}644} \approx 27{,}96$$

Concluímos que $\bar{x} = 61{,}44$, $M_e = 71{,}5$ e $s \approx 27{,}96$.

3.

3.1. A variável em estudo é o nível dos encargos, em euros, que um proprietário gasta com a compra de habitação própria. A variável é quantitativa contínua.

3.2. Com as informações da tabela, representamos o histograma de frequências absolutas seguinte.



Fonte: Censos 2021

3.3. Iremos recorrer às marcas das classes para determinar um valor aproximado da média dos níveis de encargo.

Níveis de encargos (em euros) dos proprietários com a compra de habitação própria

Nível do encargo	Frequência absoluta	Marca da classe
[0,100[	10	50
[100,200[	48	150
[200,300[	65	250
[300,400[	52	350
[400,500[	34	450
[500,600[	7	550
[600, 700[	7	650
Total	223	

Um valor aproximado para a média deste conjunto de dados é:

$$\bar{x} = \frac{10 \times 50 + 48 \times 150 + 65 \times 250 + 52 \times 350 + 34 \times 450 + 7 \times 550 + 7 \times 650}{223} \approx 295{,}29$$

Neste concelho, o nível médio de encargos gastos pelos proprietários com a compra de habitação própria é 295,29 euros.

4.

4.1. Com a representação do diagrama de extremos e quartis, concluímos que $máx. = 20$.

4.2. Este conjunto de dados tem $mín. = 5$.

4.3. O valor do primeiro quartil é $Q_1 = 10{,}5$.

4.4. Com os dados representados no diagrama de extremos e quartis, podemos afirmar que $M_e = 13{,}5$.

4.5. O valor do terceiro quartil é $Q_3 = 18{,}25$.

4.6. A amplitude deste conjunto de dados é:

- $Amplitude = máximo - mínimo = 20 - 5 = 15$

A amplitude é 15.

4.7. Para determinar a amplitude interquartil, tem-se que:

- $Amplitude\ interquartil = Q_3 - Q_1 = 18{,}25 - 10{,}5 = 7{,}75$

Neste conjunto de dados, a amplitude interquartil é 7,75.

5.

5.1. Comecemos por organizar os dados numa tabela de frequências simples.

Número de mensagens enviadas pelos alunos do 10.º A na última segunda-feira

N.º de mensagens	Frequência absoluta	Frequência relativa (%)
0	3	$\frac{3}{20} \times 100 = 15$
5	2	$\frac{2}{20} \times 100 = 10$
9	1	$\frac{1}{20} \times 100 = 5$
10	1	$\frac{1}{20} \times 100 = 5$
12	2	$\frac{2}{20} \times 100 = 10$
22	2	$\frac{2}{20} \times 100 = 10$
30	1	$\frac{1}{20} \times 100 = 5$
40	1	$\frac{1}{20} \times 100 = 5$
55	1	$\frac{1}{20} \times 100 = 5$
56	1	$\frac{1}{20} \times 100 = 5$
60	1	$\frac{1}{20} \times 100 = 5$
72	1	$\frac{1}{20} \times 100 = 5$
80	1	$\frac{1}{20} \times 100 = 5$
90	1	$\frac{1}{20} \times 100 = 5$
100	1	$\frac{1}{20} \times 100 = 5$
Total	20	100

Concluímos que a moda é 0, porque é o número de mensagens que regista a maior frequência absoluta.

5.2. O número de mensagens médio é:

$$\bar{x} = \frac{0 \times 3 + 5 \times 2 + 9 + 10 + 12 \times 2 + 22 \times 2 + 30 + 40 + 55 + 56 + 60 + 72 + 80 + 90 + 100}{20} = 34$$

O número médio de mensagens enviadas pelos alunos do 10.º A, na última segunda-feira, foi 34.

5.3. A tabela de frequências relativas simples e acumuladas deste conjunto de dados é a seguinte.

Número de mensagens enviadas pelos alunos do 10.º A na última segunda-feira

N.º de mensagens	Frequência relativa (%)	Frequência relativa acumulada (%)
0	15	15
5	10	$15 + 10 = 25$
9	5	$25 + 5 = 30$
10	5	$30 + 5 = 35$
12	10	$35 + 10 = 45$
22	10	$45 + 10 = 55$
30	5	$55 + 5 = 60$
40	5	$60 + 5 = 65$
55	5	$65 + 5 = 70$
56	5	$70 + 5 = 75$
60	5	$75 + 5 = 80$
72	5	$80 + 5 = 85$
80	5	$85 + 5 = 90$
90	5	$90 + 5 = 95$
100	5	$95 + 5 = 100$
Total	100	

Concluímos que $P_{60} = 30$. Tal significa que, pelo menos, 60% dos alunos da turma enviaram até 30 mensagens de WhatsApp.

5.4. Com os dados ordenados, por ordem crescente, podemos verificar que os extremos e quartis deste conjunto de dados são os seguintes.

$Mín. = 0$ $\qquad$ $Máx. = 100$

0	0	0	5	5	9	10	12	12	22	22	30	40	55	56	60	72	80	90	100

$$M_e = \frac{22 + 22}{2} = 22$$

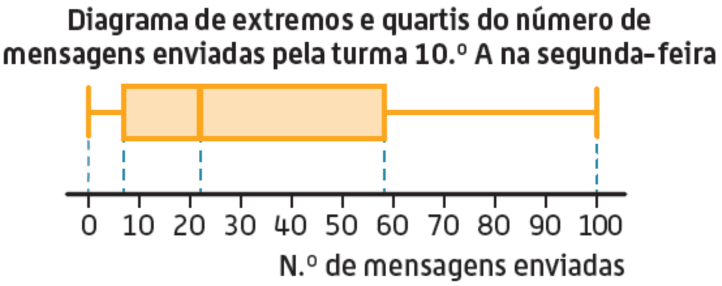
0	0	0	5	5	9	10	12	12	22

$$Q_1 = \frac{5 + 9}{2} = 7$$

22	30	40	55	56	60	72	80	90	100

$$Q_3 = \frac{56 + 60}{2} = 58$$

Podemos então construir o diagrama de extremos e quartis seguinte.



Fonte: Turma do 10.º A

6. Sabemos que a média do número de vezes que os alunos da turma da Mariana frequentaram o refeitório é igual a 20.

Assim:

$$\bar{x} = 20 \Leftrightarrow \frac{0 \times 3 + 1 \times 2 + 5 \times 2 + 6 \times 2 + 10 \times 1 + 12 \times 2 + 14 \times 1 + 16 \times a + 20 \times 4}{3 + 2 + 2 + 2 + 1 + 2 + 1 + a + 4} = 20 \Leftrightarrow$$

$$\Leftrightarrow \frac{0 + 2 + 10 + 12 + 10 + 24 + 14 + 16a + 80}{3 + 2 + 2 + 2 + 1 + 2 + 1 + a + 4} = 20 \Leftrightarrow \frac{152 + 16a}{17 + a} = 20 \Leftrightarrow 152 + 16a = 20 \times (17 + a) \Leftrightarrow$$

$$\Leftrightarrow 152 + 16a = 140 + 20a \Leftrightarrow 16a - 20a = 140 - 152 \Leftrightarrow -4a = -12 \Leftrightarrow a = 3$$

O total de alunos inquiridos é:

$$total = 3 + 2 + 2 + 2 + 1 + 2 + 1 + 3 + 4 = 20$$

A mediana corresponde à média aritmética das respostas dos alunos que ocupam a posição 10 e a posição 11, no caso de ordenarmos os dados por ordem crescente. Ou seja:

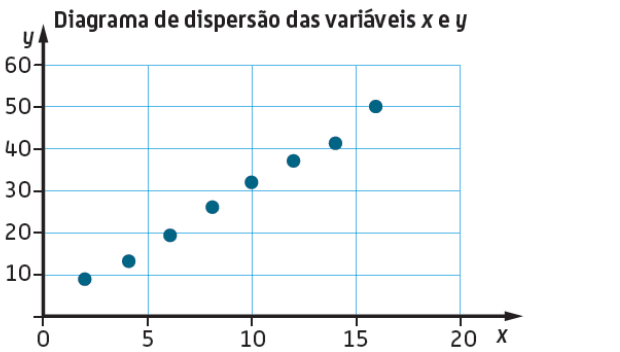
0	0	0	1	1	5	5	6	6	10	12	12	14	16	16	16	20	20	20	20

$$M_e = \frac{x_{10} + x_{11}}{2} = \frac{10 + 12}{2} = 11$$

A mediana deste conjunto de dados é 11.

7.

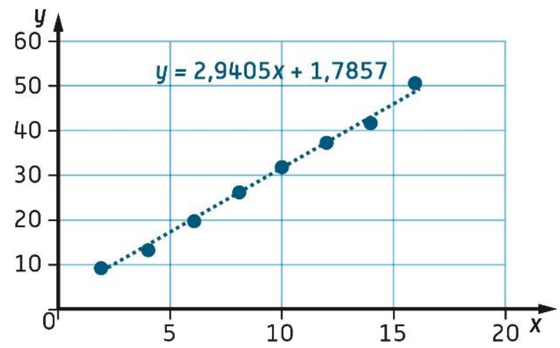
7.1. Seja o conjunto de pares ordenados (x_i, y_i). Com este conjunto, conseguimos representar o diagrama de dispersão seguinte.



Fonte: Própria

7.2. Considera a equação da reta de regressão $y = a + b$, com a e b constantes reais.
Recorrendo às potencialidades de uma folha de cálculo, obtemos os valores seguintes.

Diagrama de dispersão e reta de regressão das variáveis x e y



Fonte: Própria

$$a \approx 2{,}940 \qquad b \approx 1{,}786$$
$$y = 2{,}9405x + 1{,}7857$$

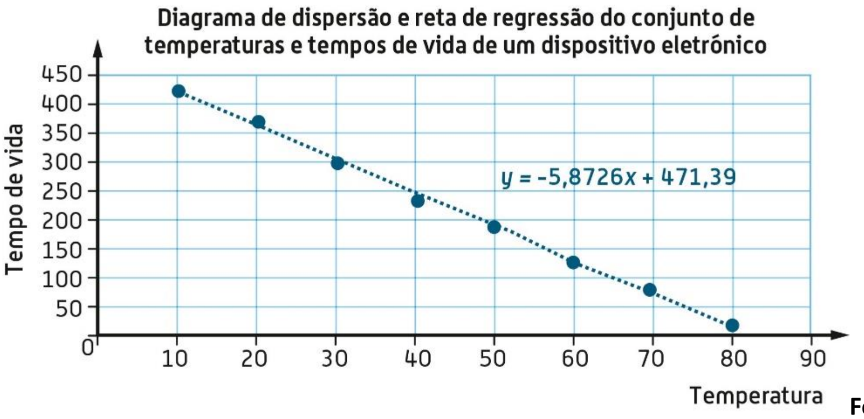
A equação da reta de regressão é $y = 2{,}940x + 1{,}786$.

7.3. Usando a folha de cálculo, o coeficiente de correlação é:

$$r \approx 0{,}998$$

Trata-se de uma correlação positiva forte, pois o coeficiente de correlação apresenta um valor muito próximo de 1.

8. Utilizando as potencialidades de uma folha de cálculo, começamos por representar o diagrama de dispersão e a reta de regressão seguintes.



Fonte: Empresa

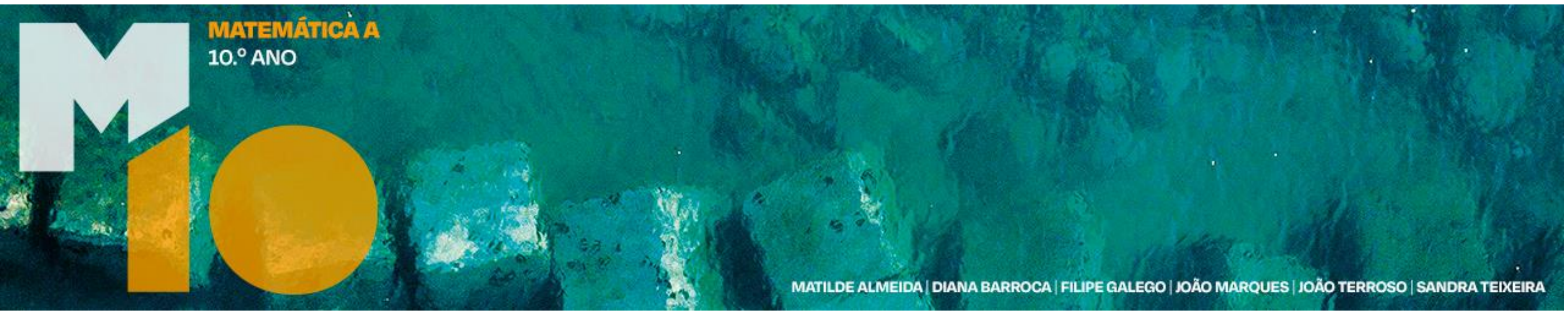
A equação da reta de regressão, obtida na folha de cálculo, é:

$$y = -5{,}8726x + 471{,}39$$

O tempo de vida previsível de um equipamento, quando sujeito a uma temperatura de 35 graus Celsius ($x = 35$) é:

$$y = -5{,}8726 \times 35 + 471{,}39 = 265{,}849 \approx 266$$

Um equipamento de fica sujeito a uma temperatura de 35 graus Celsius tem como previsão de vida útil 266 dias, aproximadamente.



Preparar o Exame - Página 182 a 185

1.

1.1. A afluência às urnas tem por base o nível de abstenção registado nesta eleição. A taxa de abstenção é determinada da forma seguinte.

- $abstencionistas = 569 - (132 + 132 + 69 + 68 + 100 + 22) = 46$
- $taxa\ de\ abstenção = \frac{46}{569} \times 100 \approx 8{,}08\ \%$

Como a taxa de abstenção é de, aproximadamente, 8%, podemos concluir que a afluência às urnas de voto foi elevada, daí o comentário de contentamento do Presidente da Assembleia.

1.2. Fazendo a recolha das pontuações de cada uma das medidas, tal como prevista no Regimento Interno desta assembleia, constatamos que:

$Medida\ A \rightarrow 4 \times 132 + 3 \times 69 + 2 \times 68 + 1 \times (100 + 22) = 993\ pontos$
$Medida\ B \rightarrow 4 \times 22 + 3 \times 100 + 2 \times 69 + 1 \times (132 + 68) = 726\ pontos$
$Medida\ C \rightarrow 4 \times 69 + 3 \times (132 + 68) + 2 \times (100 + 22) + 1 \times 0 = 1120\ pontos$
$Medida\ D \rightarrow 4 \times (100 + 68) + 3 \times 22 + 2 \times 132 + 1 \times 69 = 1071\ pontos$

Segundo o Regimento Interno, a medida C é a medida vencedora.

R.: Opção (C)

2.

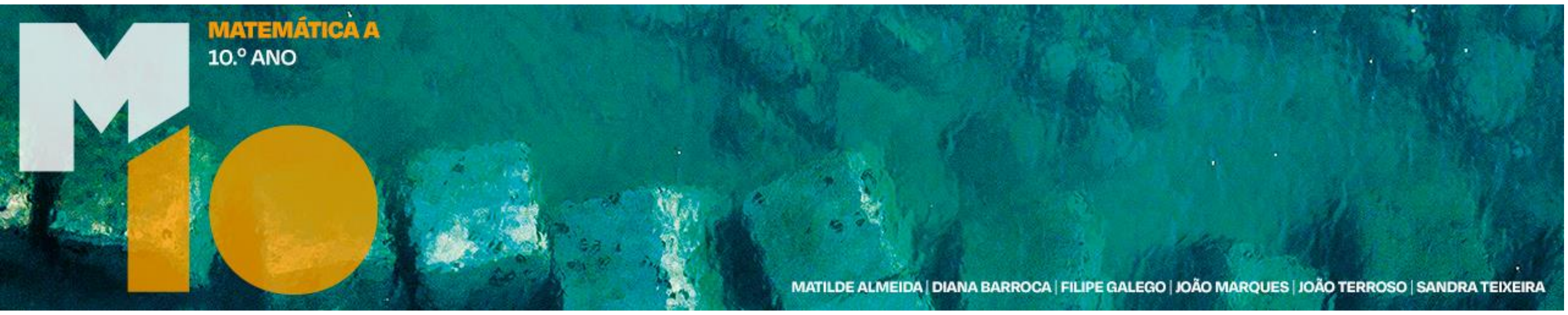
2.1. Como previsto pelo Método de Hondt, comecemos por determinar os quocientes resultantes da divisão do número de votos validamente expressos de cada um dos partidos pelos divisores 1, 2, 3, 4, e assim sucessivamente.

Partido	PSOO	PPE	CS	LM	EV	JNT
Número de VVE / Divisores	33 180	20 350	12 300	10 170	6280	5640
1	33 180	20 350	12 300	10 170	6280	5640
2	16 590	10 180	6150	5085	3140	2820
3	11 060	6783	4100	3390	2093	1880
4	8295	5088	3075	2543	1570	1410
5	6636	4070	2460	2034	1256	1128
6	5530	3392	2050	1695	1047	940

Logo, atribuímos um mandato a cada um dos 12 maiores quocientes registados desta eleição.

Partido	PSOO	PPE	CS	LM	EV	JNT
Número de VVE / Divisores	33 180	20 350	12 300	10 170	6280	5640
1	33 180 (1)	20 350 (2)	12 300 (4)	10 170 (7)	6280 (11)	5640
2	16 590 (3)	10 180 (6)	6150 (12)	5085	3140	2820
3	11 060 (5)	6783 (9)	4100	3390	2093	1880
4	8295 (8)	5088	3075	2543	1570	1410
5	6636 (10)	4070	2460	2034	1256	1128
6	5530	3392	2050	1695	1047	940

PSOO receberia 5 mandatos. PPE receberia 3 mandatos. CS receberia 2 mandatos. LM receberia 1 mandato. EV receberia 1 mandato.



2.2.

a) Tal como prevista pela Lei Eleitoral, comecemos por determinar os quocientes obtidos pela divisão do número de votos validamente expressos de cada uma das forças políticas por 1, 3, 5, 7, 9, e assim sucessivamente.

Partido	PSOO	PPE	CS	LM	EV	JNT
Número de VVE / Divisores	33 180	20 350	12 300	10 170	6280	5640
1	33 180	20 350	12 300	10 170	6280	5640
3	11 060	6783	4100	3390	2093	1880
5	6636	4070	2460	2034	1256	1128
7	4740	2907	1757	1453	897	806
9	3687	2261	1367	1130	698	627
11	3016	1850	1118	925	571	513

Os mandatos são distribuídos pelas diferentes forças políticas, tendo em conta os 12 maiores quocientes registados.

Partido	PSOO	PPE	CS	LM	EV	JNT
Número de VVE / Divisores	33 180	20 350	12 300	10 170	6280	5640
1	33 180 (1)	20 350 (2)	12 300 (3)	10 170 (5)	6280 (8)	5640 (9)
3	11 060 (4)	6783 (6)	4100 (11)	3390	2093	1880
5	6636 (7)	4070 (12)	2460	2034	1256	1128
7	4740 (10)	2907	1757	1453	897	806
9	3687	2261	1367	1130	698	627
11	3016	1850	1118	925	571	513

Concluímos então os mandatos foram distribuídos da seguinte forma: 4 mandatos para o PSOO; 3 mandatos para o PPE; 2 mandatos para o CS; 1 mandato para cada uma das restantes, LM, EV e JNT.

b) Considerando os quocientes obtidos pela aplicação do método descrito na Lei Eleitoral de Portulália, verificamos que PSOO será a próxima força política a obter mais um mandato, pela ordem decrescente dos quocientes.

Partido	PSOO	PPE	CS	LM	EV	JNT
Número de VVE / Divisores	33 180	20 350	12 300	10 170	6280	5640
1	33 180 (1)	20 350 (2)	12 300 (3)	10 170 (5)	6280 (8)	5640 (9)
3	11 060 (4)	6783 (6)	4100 (11)	3390	2093	1880
5	6636 (7)	4070 (12)	2460	2034	1256	1128
7	4740 (10)	2907	1757	1453	897	806
9	3687 (13)	2261	1367	1130	698	627
11	3016	1850	1118	925	571	513

Assim, será necessário a atribuição de 13 mandatos para que PSOO receba mais um representante na Assembleia Nacional.
Opção (A)

3.

3.1. A base salarial bruta por hora recebida pela Carolina é obtida tendo por base a seguinte fórmula.

- $salário\ por\ hora = \frac{salário\ bruto\ anual}{52\times 40}$

Logo:

- $7{,}02 = \frac{salário\ bruto\ anual}{52\times 40} \Leftrightarrow salário\ bruto\ anual = 7{,}02 \times 52 \times 40 \Leftrightarrow salário\ bruto\ anual = 14\,601{,}60$ €

Opção (B)

3.2. A remuneração bruta mensal da Carolina é dada por:

$$salário\ bruto\ mensal = \frac{14\,601{,}60}{14} \approx 1042{,}97 \text{ €}$$

Logo, a Carolina irá reter em fonte de IRS 11,2% do seu rendimento bruto mensal. Ou seja:

$$retenção\ em\ fonte\ de\ IRS\ anual = 1042{,}97 \times 0{,}112 \times 12 \approx 1401{,}75 \text{ €}$$

Anualmente, a Carolina retém em fonte de IRS 1401,75 euros, aproximadamente.

3.3. Tendo em conta as condições do depósito realizado pela Carolina, temos que o capital acumulado por esta trabalhadora é:

- $C = 2\,340$ € $\quad i = 0{,}03 \quad k = 4$
- $C_5 = 2340 \times \left(1 + \frac{0{,}03}{4}\right)^{4\times 5} \approx 2717{,}17$ €

Logo:

- $lucro = 2\,717{,}17 - 2\,340 = 377{,}17$ €

O lucro que a Carolina irá obter com este depósito é de 377,17 euros.

4.

4.1. Para completar a tabela, deveremos determinar a diferença entre a hora de chegada e a hora prevista de chegada deste conjunto de 10 voos da forma seguinte.

Hora prevista	Hora de Chegada	Tempo do atraso (em minutos)
15:40	16:18	$20 + 18 = 38$
15:40	15:58	18
15:45	16:10	$15 + 10 = 25$
16:00	16:50	50
16:00	16:05	5
16:00	16:21	21
16:05	16:28	23
16:10	16:15	5
16:15	16:25	10
16:20	17:07	$40 + 7 = 47$

4.2. Considerando a variável tempo do atraso (em minutos), podemos representar os dados no diagrama de caule-e-folhas seguinte.

Diagrama de caule-e-folhas dos tempos, em minutos, de atraso de 10 voos do Aeroporto Humberto Delgado

0	5	5	
1	0	8	
2	1	3	5
3	8		
4	7		
5	0		

Legenda: 0|5 representa 5 minutos de atraso

Fonte: Própria

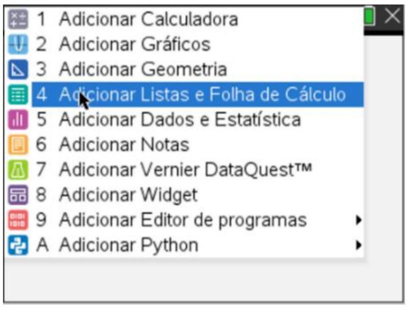
4.3. A média dos atrasos dos 10 voos considerados é:

$$\bar{x} = \frac{38+18+25+50+5+21+23+5+10+47}{10} = \frac{242}{10} = 24{,}2$$

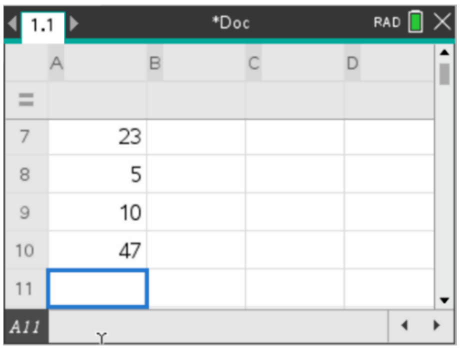
A média dos atrasos deste conjunto de voos é 24,2 minutos.

4.4.

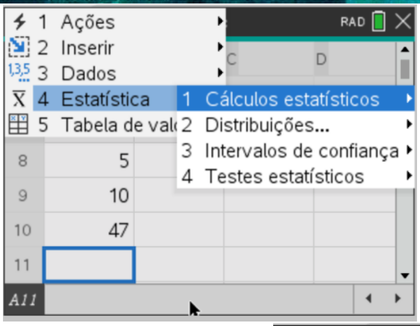
Recorrendo às potencialidades da calculadora gráfica, começamos por recorrer ao *software* folha de cálculo.

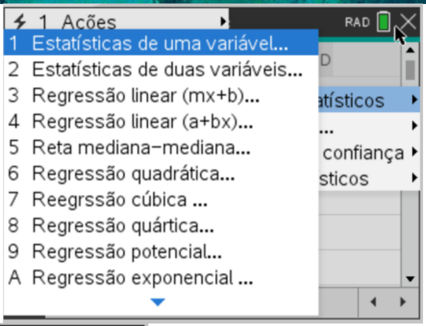


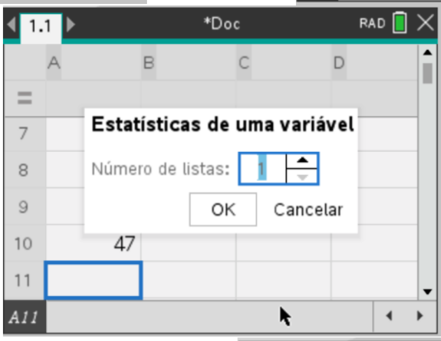
De seguida, transcreve-se para uma lista da folha de cálculo os dados estatísticos associados à variável tempo de atraso (em minutos):



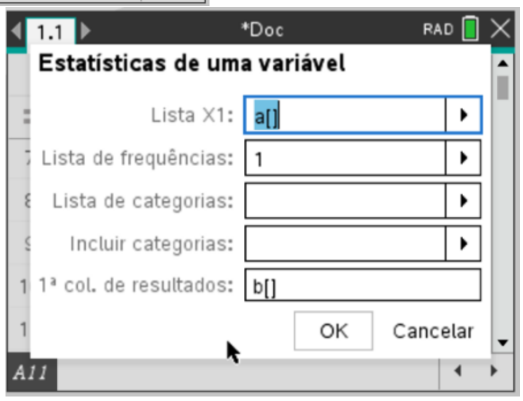
Selecionando em *menu* , *Estatística* a opção *1 Cálculos estatísticos* e, de seguida, *1 Estatísticas de uma variável*, identificamos o número de lista com 1.



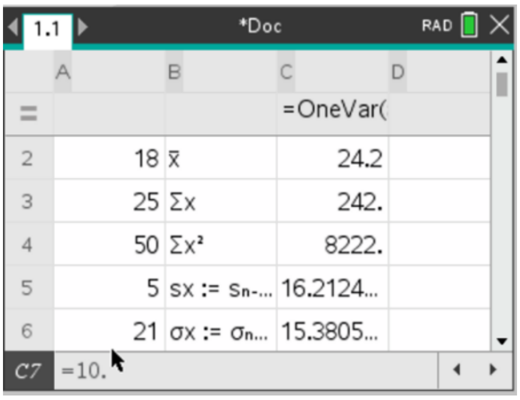




Posteriormente, seleciona-se a lista e considera-se a *Lista de frequência* como 1.



Depois destes procedimentos, obtém-se os seguintes resultados:



Ao consultar as informações fornecidas pela calculadora, conseguimos concluir que o valor do desvio padrão é de 16,2, aproximadamente.

5. Para determinar um valor aproximado da média das massas das mochilas iremos utilizar as marcas das classes consideradas.

Classe	Marca da Classe
$[0,1[$	$\frac{0+1}{2}=0{,}5$
$[1,2[$	$\frac{1+2}{2}=1{,}5$
$[2,3[$	$\frac{2+3}{2}=2{,}5$
$[3,4[$	$\frac{3+4}{2}=3{,}5$

A média é determinada por: $\bar{x}=\frac{0{,}5\times5+1{,}5\times6+2{,}5\times10+3{,}5\times3}{5+6+10+3}\approx1{,}96$

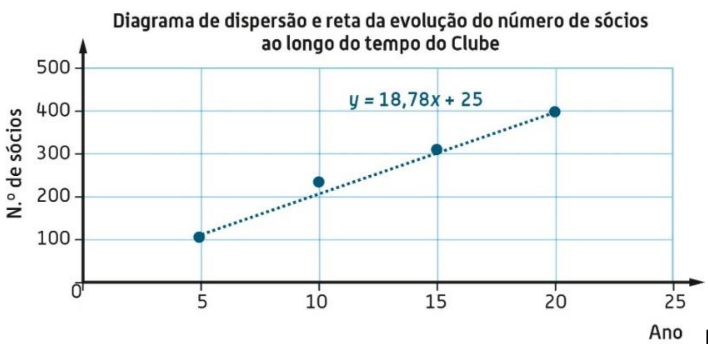
As massas das mochilas dos alunos desta escola pesam, em média, $1{,}96$ kg, aproximadamente.

6. Observando o diagrama de dispersão, podemos deduzir que a reta de regressão tem declive positivo e ordenada na origem negativa. Assim, poderemos considerar as opções (A) e (D) como corretas.
No entanto, o coeficiente de correlação tem de ser positivo, pelo que a resposta correta é a opção (D).

7. Comecemos por considerar que $t=0$ corresponde ao ano 2000.
Usando as potencialidades de uma folha de cálculo, introduzimos as listas seguintes.

Lista 1	Lista 2
5	105
10	232
15	310
20	392

De seguida, construímos o diagrama de dispersão e a reta de regressão, obtendo a representação gráfica seguinte.



Fonte: CFNL

Com a equação da reta de regressão obtida, estimamos o número de sócios (y) no final de 2022, considerando $x=22$:

$$y=18{,}8\times22+25\approx438$$

No final de 2022, estima-se que o CFNL tenha 438 sócios, aproximadamente.